Transcription of eukaryotic genome

Transcription of eukaryotic genome

- **Primary transcripts**
  - **precursor messenger mRNA (pre-mRNA)**
  - **heterogeneous nuclear RNA (hnRNA) = pre-mRNA forming in nucleus**
  - **precursor ribosomal RNA (pre-rRNA)**
  - **precursor transfer RNA (pre-tRNA)**
  - **5S-rRNA**
  - **small RNA (snRNA, snoRNA, scRNA)**

- **Eukaryotic DNA-dependent RNA-polymerase**

  - RNA-polymerase I, II, III

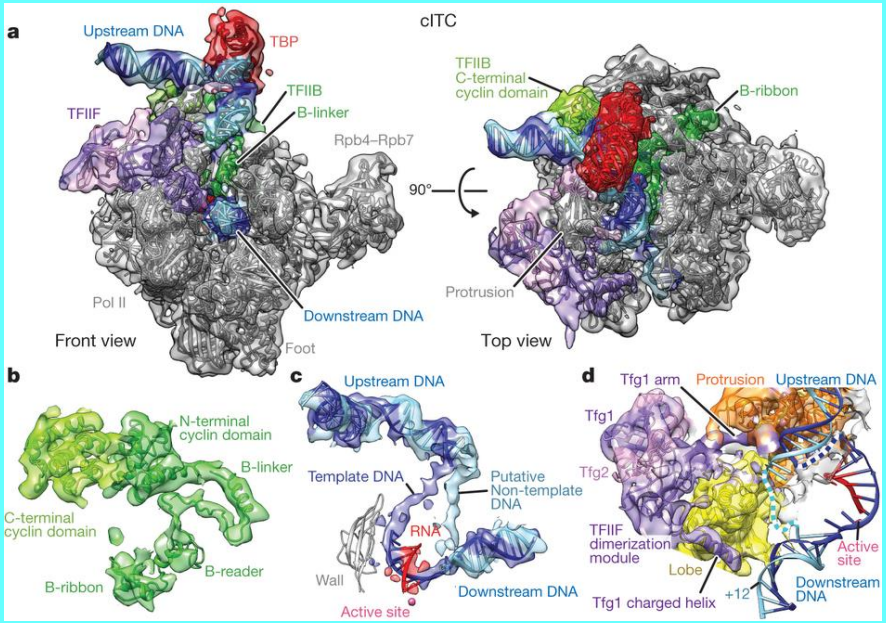
- **Transkription factors**

Eukaryotic RNA-polymerases

- RNA polymerase I
  - Synthesis of pre-rRNA
  - Only in nucleolus
  - Not sensitive to α-amanitin
- **RNA polymerase II**
  - **Synthesis of hnRNA and some snRNA**
  - **Sensitive to α-amanitin**
- RNA polymerase III
  - Synthesis of pre-tRNA, 5S-rRNA and some snRNA

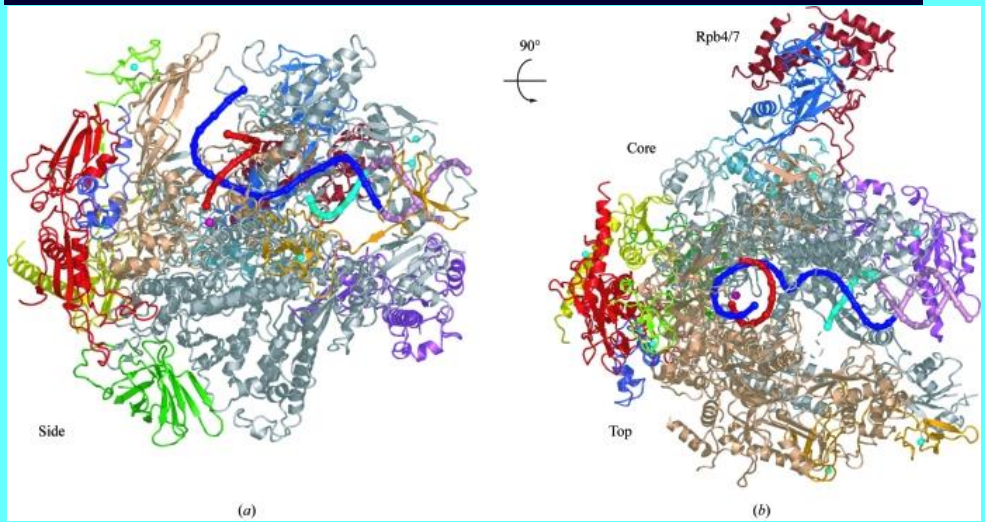
RNA polymerase II

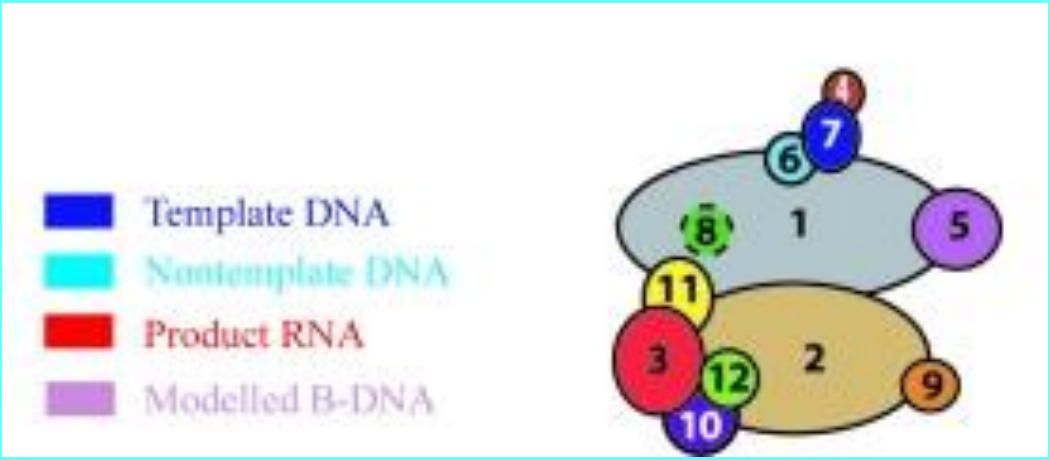
- **Compound from 12 subunits**



Nature **518**, 376–380 (19 February 2015) doi:10.1038/nature14229

Acta Crystallogr D Biol Crystallogr. 2009 Feb 1; 65(Pt 2):112-120





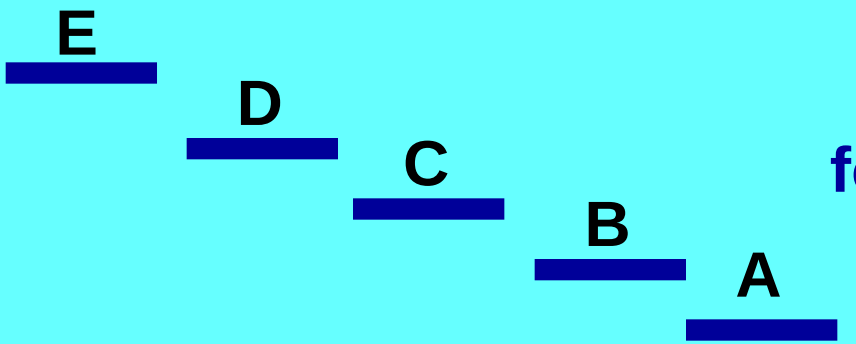
Transcription unit of prokaryotes

polycistronic character

***trp* operon in *Escherichia coli* – 5 genes**

***trp* mRNA**

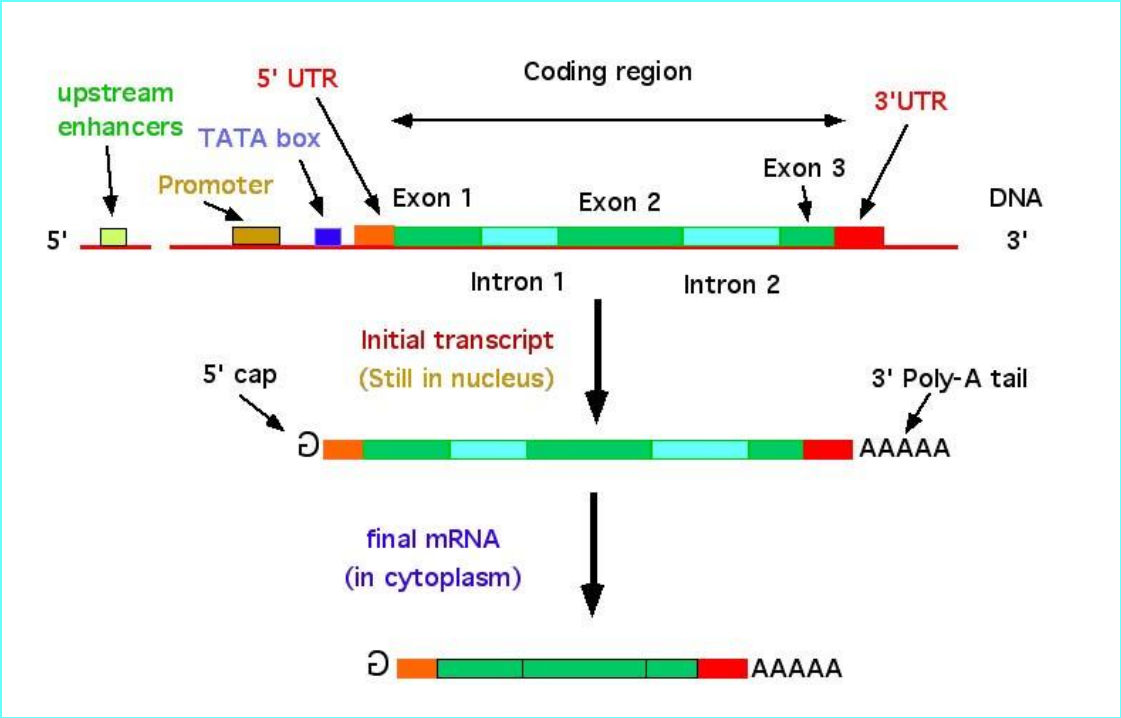
origins of translation

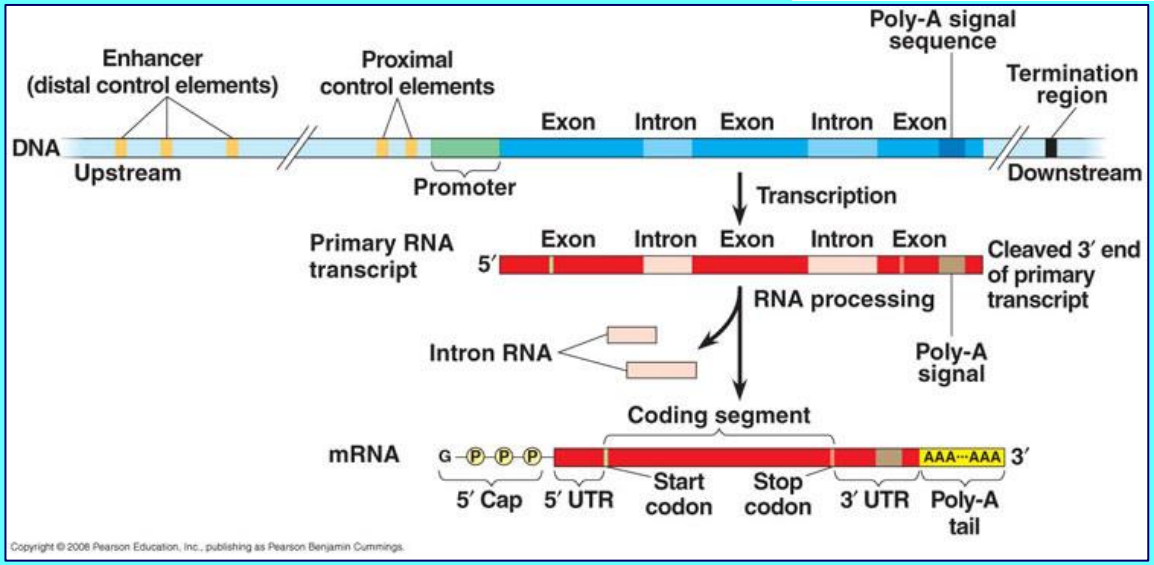


5 proteins for synthesis of Trp

Transcription unit of eukaryotes

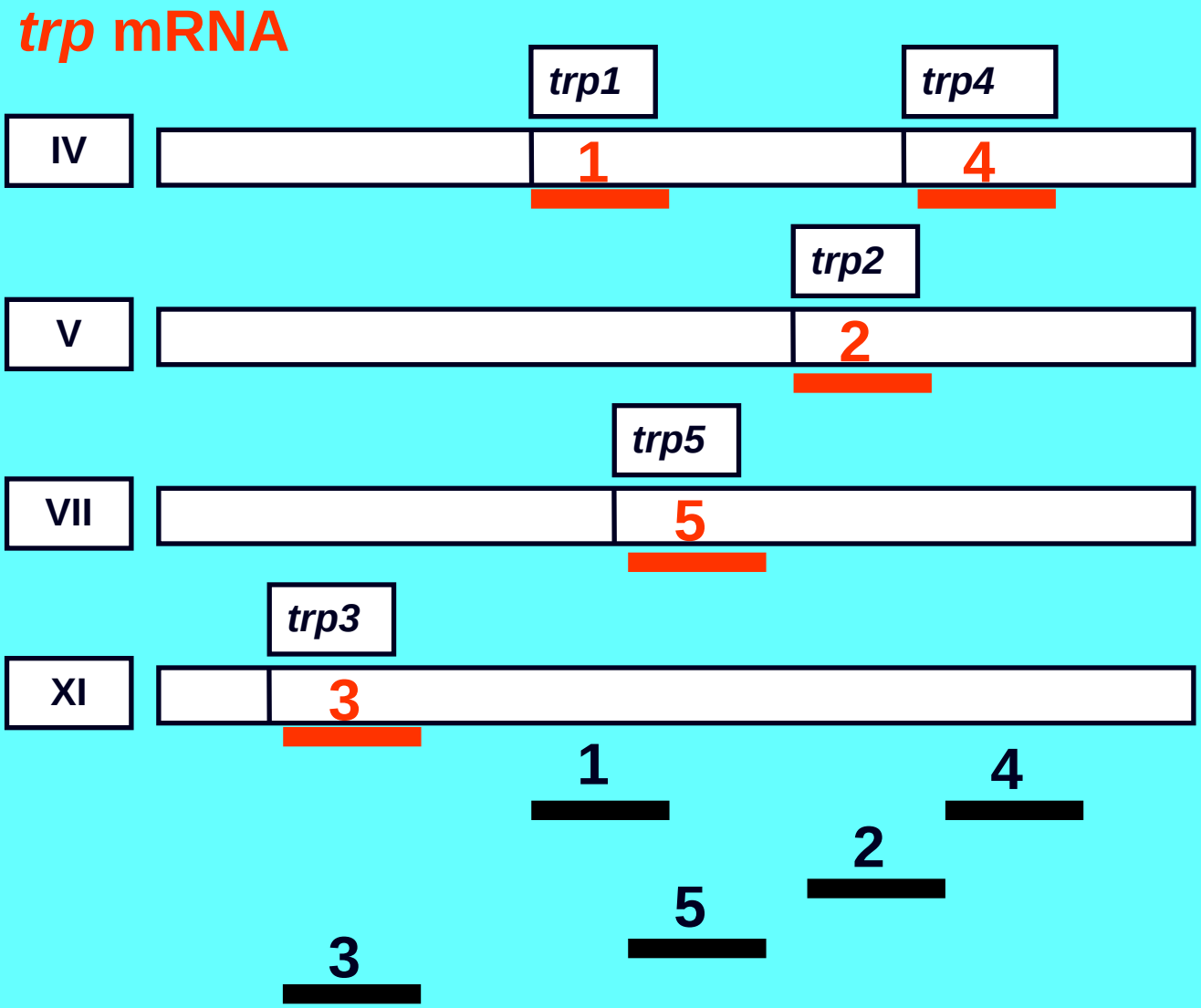
- Monocistroni character
- Contains:
  - Promoter
  - Leading sequence (5´-UTR)
  - Polyadenilation signal
  - Terminator





http://nitro.biosci.arizona.edu/courses/EEB600A-2003/lectures/lecture24/lecture24.html

The transcription unit of S. cerevisiae



Transcription unit for synthesis of Trp in *Saccharomyces cerevisiae*

= togehter 5 genes located on 4 chromosomes

Transcription factors

- **Regulatory elements necessary for transcription initiation**
- **Usually initiate transcription, rarely inhibit it**
- **Their different combination bind to the promoter, then the RNA polymerase bind to DNA strand**

Types of transcription factors

- **general TF**
  - **present in all and most types of cells**
  - **necessary to transcription initiation**
  - **basal – low activity, minimal cell requirements**
  - **constitutive – increase the basal activity according to cell type; basal cell requirements**
- **special TF**
  - **only in cells of specific tissues and in a certain time**
  - **applied in inducible transcription**

Promoter of RNA-polymerase II

PROKARYOTIC PROMOTER

starting nucleotide +1

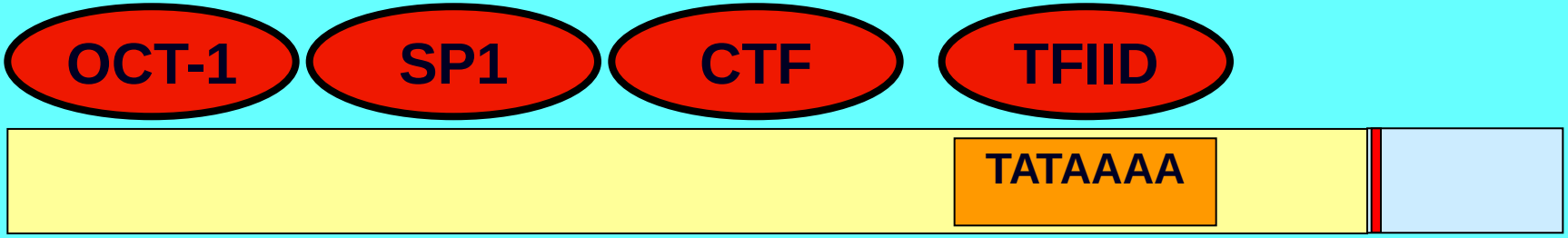
region -35

Pribnow box
region -10

EUKARYOTIC PROMOTER

constitutive and special transcription factors

basal transcription factors

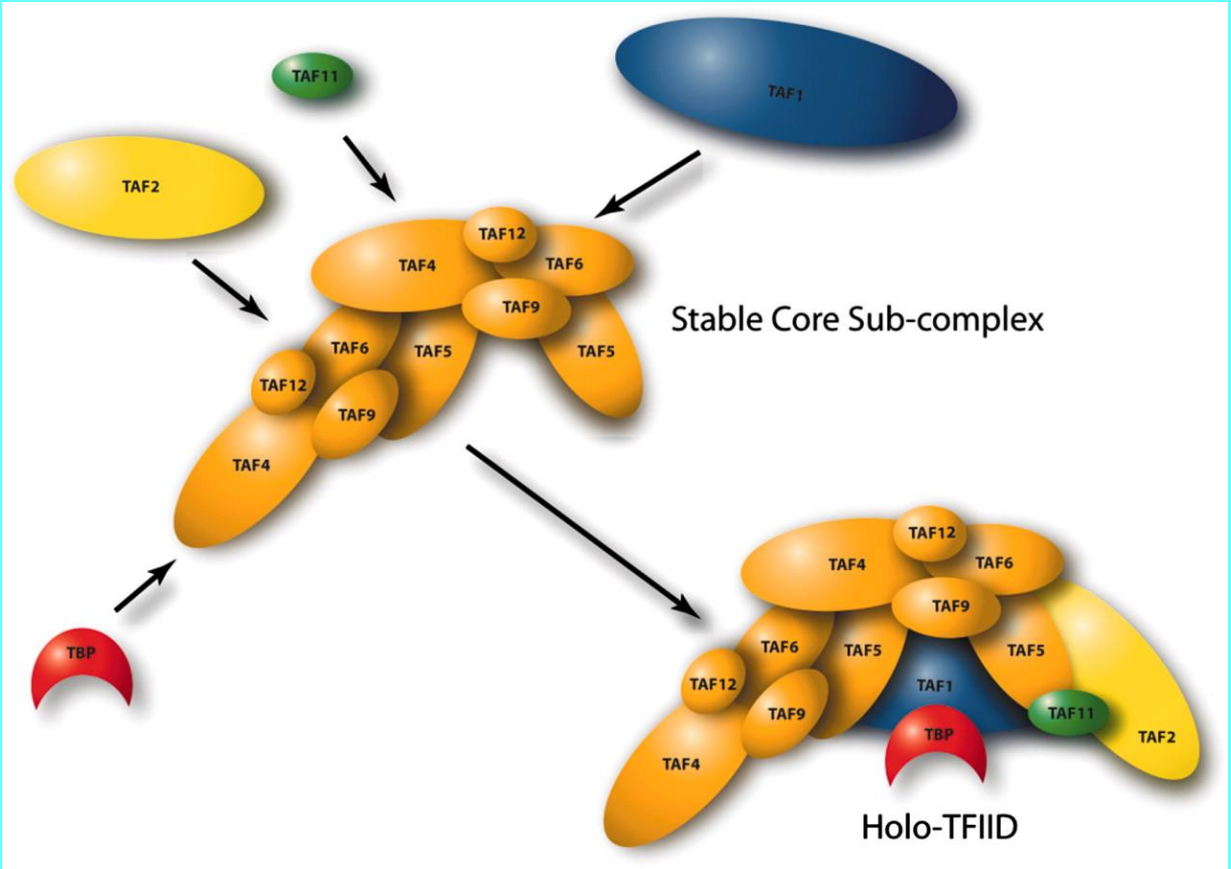


Hogness box (TATA box)
region -34 to -26

Binding to TATA box

1) Recognises by basal TF TFIID
2) Part of TFIID is TBP protein (TATA binding protein), which is present in all eukyotes

Model of TFIID assembly *in vivo*



Kevin J. Wright et al. PNAS 2006;103:12347-12352

Transcription of hnRNA

1) Separation of transcription and translation
2) hnRNA is capped by the cap, and methylated (binding to ribosome)
3) In the 3´- region (after STOP codon) the sequence AAUAAA is present, in this location the hnRNA is digested
4) At 3´end is polyadenylated (stabilisation in cytoplasm)
5) After removing introns and joining exons it is transformed to mRNA

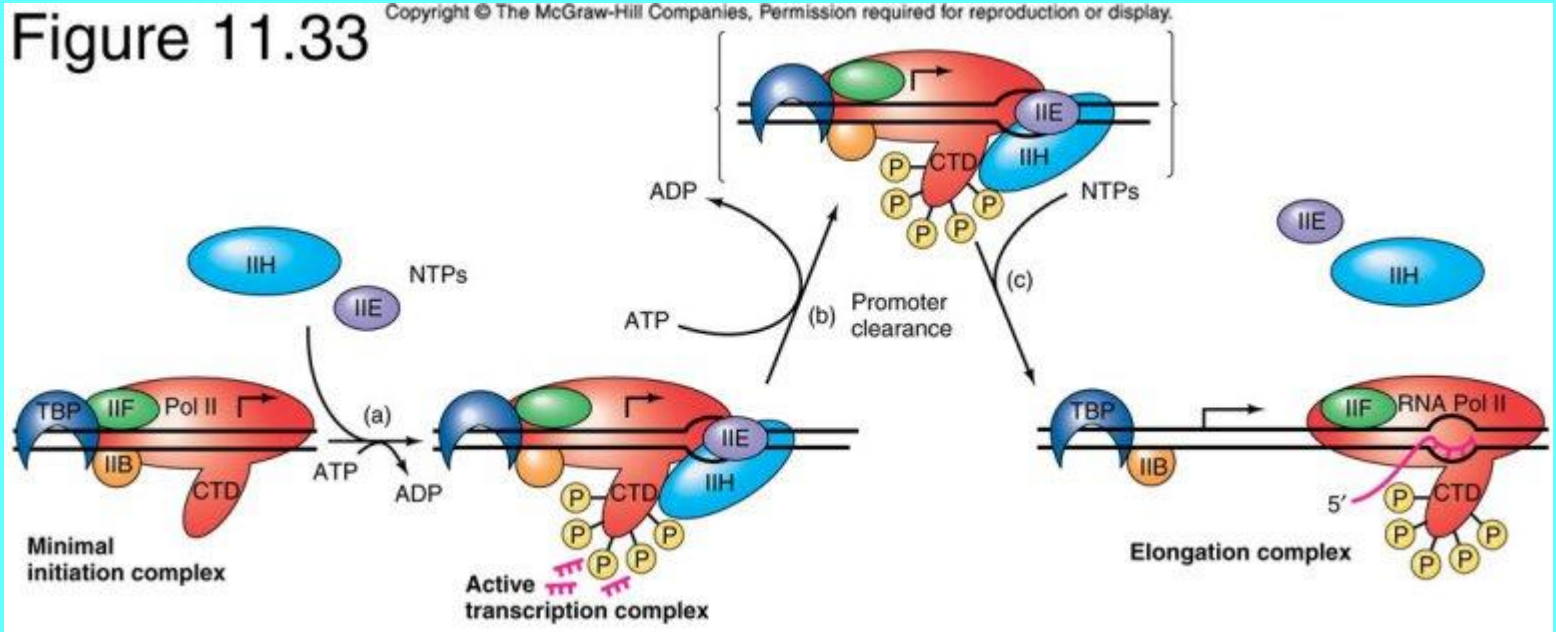
Initiation of transcription

1) Binding of transcription factors on TATA box and others regulation sequences = preinitiation complex

2) Binding of RNAP II on preinitiation complex = closed initiation complex

3) Phosphorylation of CTD domain of RNAP II by trancription factor TFIIH (halicase and kinase activities) → RNAP II activation and unwinding of dsDNA = open initiation complex

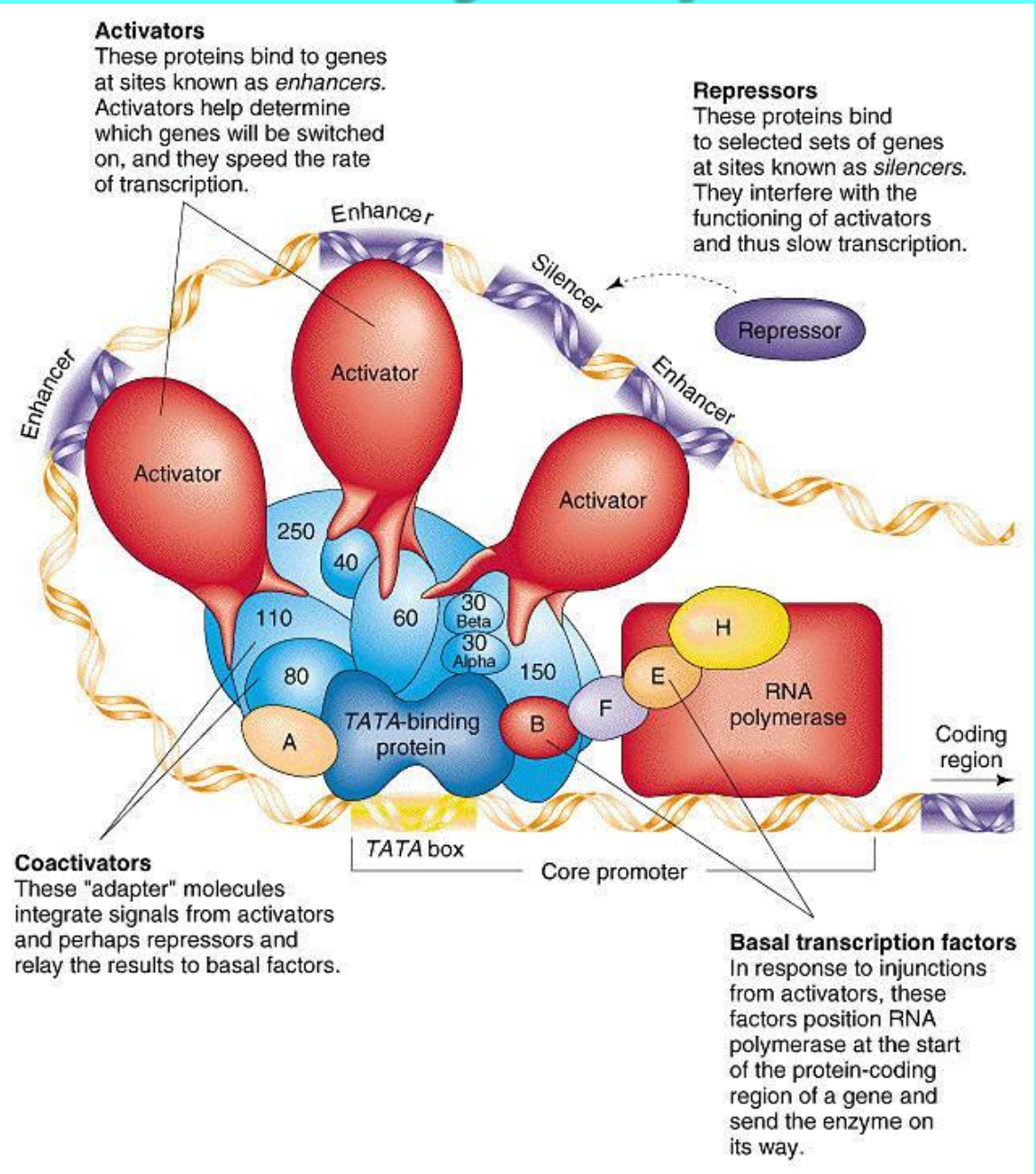
4) Disociation of RNAP II from TFs (except TFIIF) and start of RNA synthesis



Eukaryotic transcription - video

https://www.youtube.com/watch?v=icZjgZozkB8

Parts of eukaryotic promoter

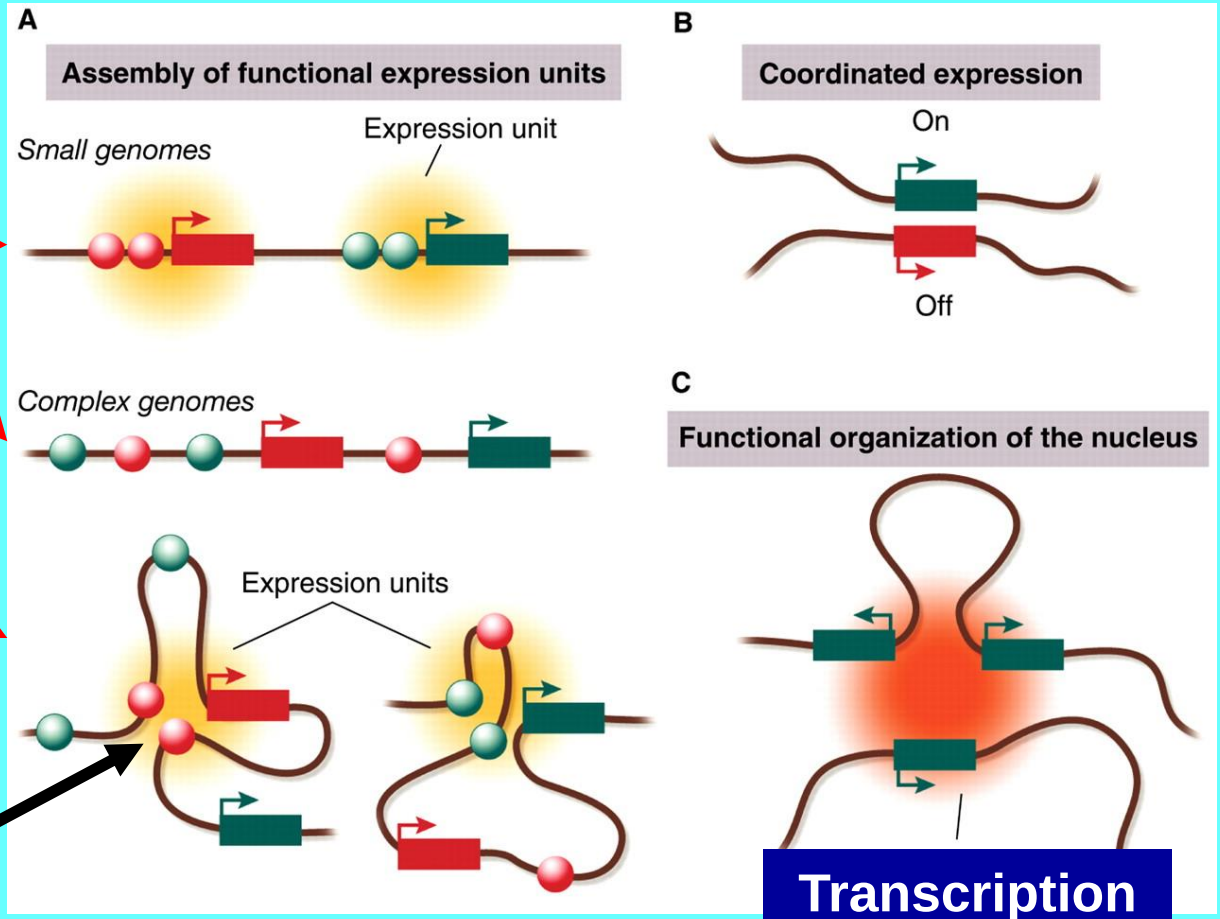


http://www.cbs.dtu.dk/dtucourse/cookbooks/dave/Lekt03bkg.html

Spatial assemblies of transcription

Linear transcription units

Spatial transcription units



Promoter assembled from 2 promoters?

Transcription factory

Dekker J.: Science 319, 1793 -1794 (2008)

Original text to previous picture

Spatial assemblies

(A) Linearly defined expression units in compact genomes and spatially assembled expression units in complex genomes.
(B) Association between coordinately expressed genes.
(C) Colocalization of genes at subnuclear structures, such as transcription factories.

Circles, regulatory elements; rectangles, genes. Arrows indicate direction of transcription.

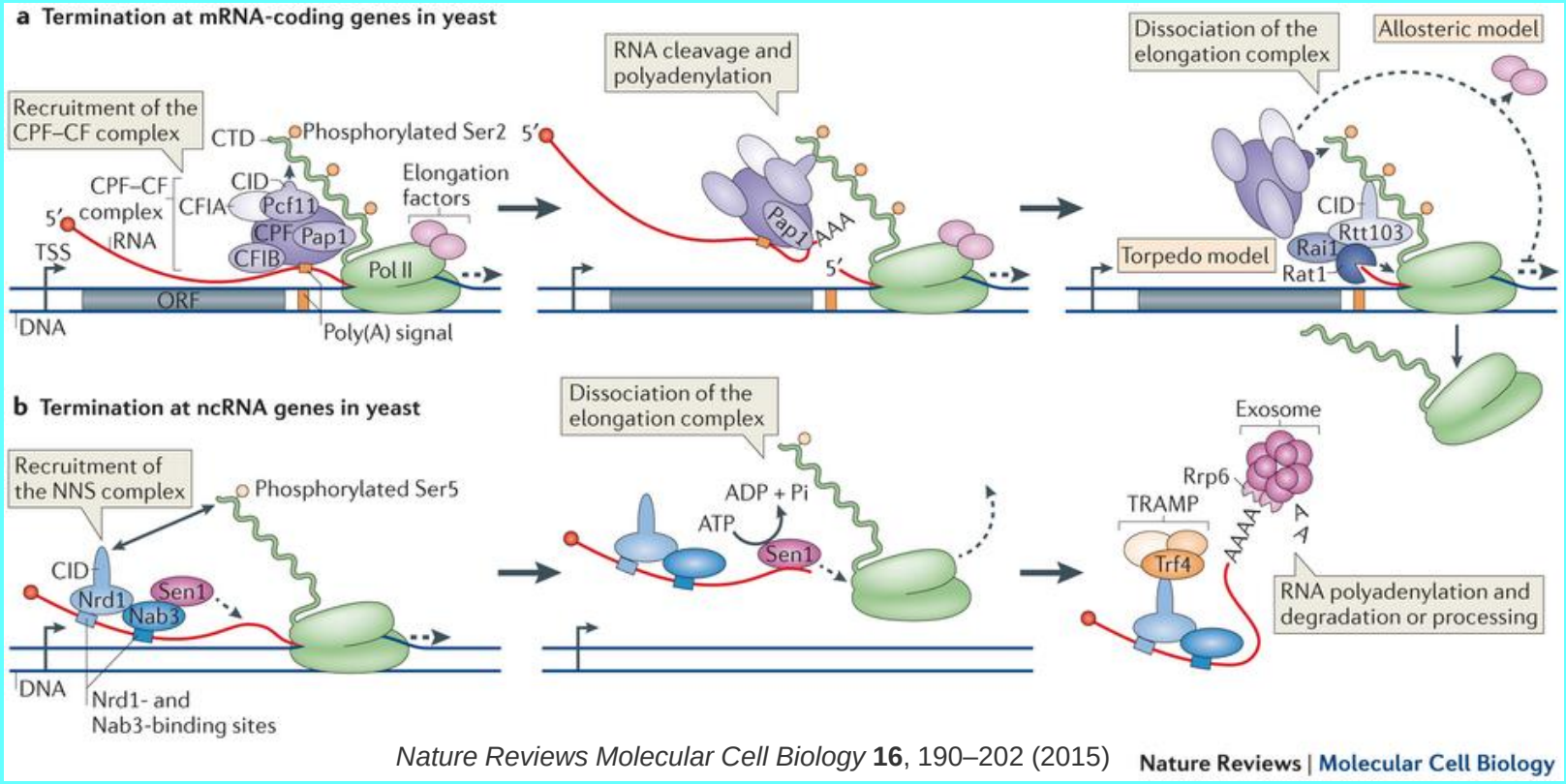
Dekker J.: Science 319, 1793 -1794 (2008)

Termination of transcription

1) Terminator contains AATAAA sequence = <u>polyadenilation signal</u>

2) Once polyadenilation signal is transcripted into hnRNA, it is recognised by protein complex, which cut hnRNA 10-30 nt towards 3´-end

3) Subsequently, RNAP II disociate from DNA and the rest of hnRNA behind the polyadenilation signal is degraded



Nature Reviews Molecular Cell Biology **16**, 190–202 (2015) Nature Reviews | Molecular Cell Biology

Transcription and nucleoporins

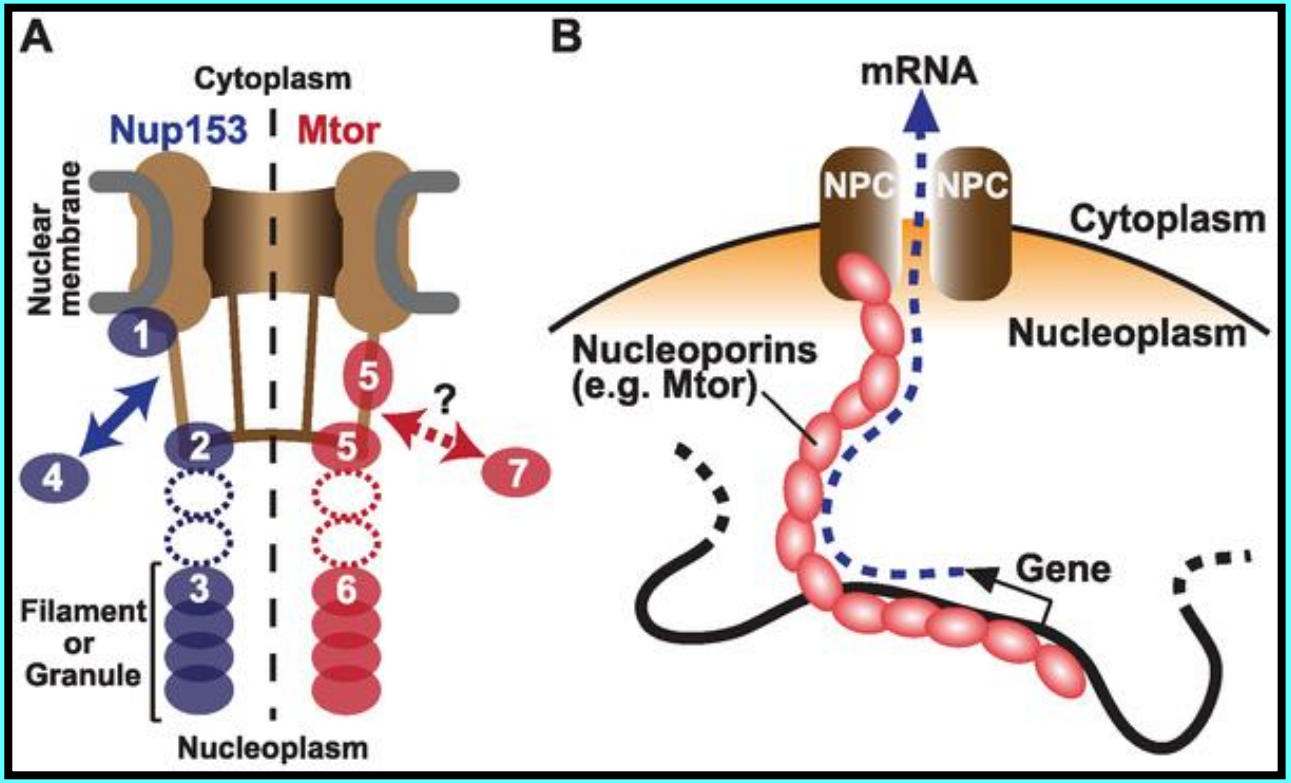
- In yeast, frequently transcribed genes are localised near by nuclear pores
- Also after transcription activation activated regions are transported from central part of nucleus to its surface
- Yeast have no lamins which are localised in inner surface of nucleolema
- Multicellular organisms have lamins

Lamins bind to heterochromatin, they deactivate the gene expression

Ikegami, K. a Lieb, J. D. PlosGenetics 6 (2), 1-2 (February 2010)

Transcription and nucleoporins

- Nuclear pore complex (NPC) selectively transmit macromolecules
- They are complexes of more than 400 proteins (nucleoporins) which create about 30 subunits



Nucleoporins Nup153 and Mtor form filamentous structures which transport DNA from inner part of nucleus to nuclear pores

Ikegami, K. a Lieb, J. D. PlosGenetics 6 (2), 1-2 (February 2010)

Posttranscription RNA processing

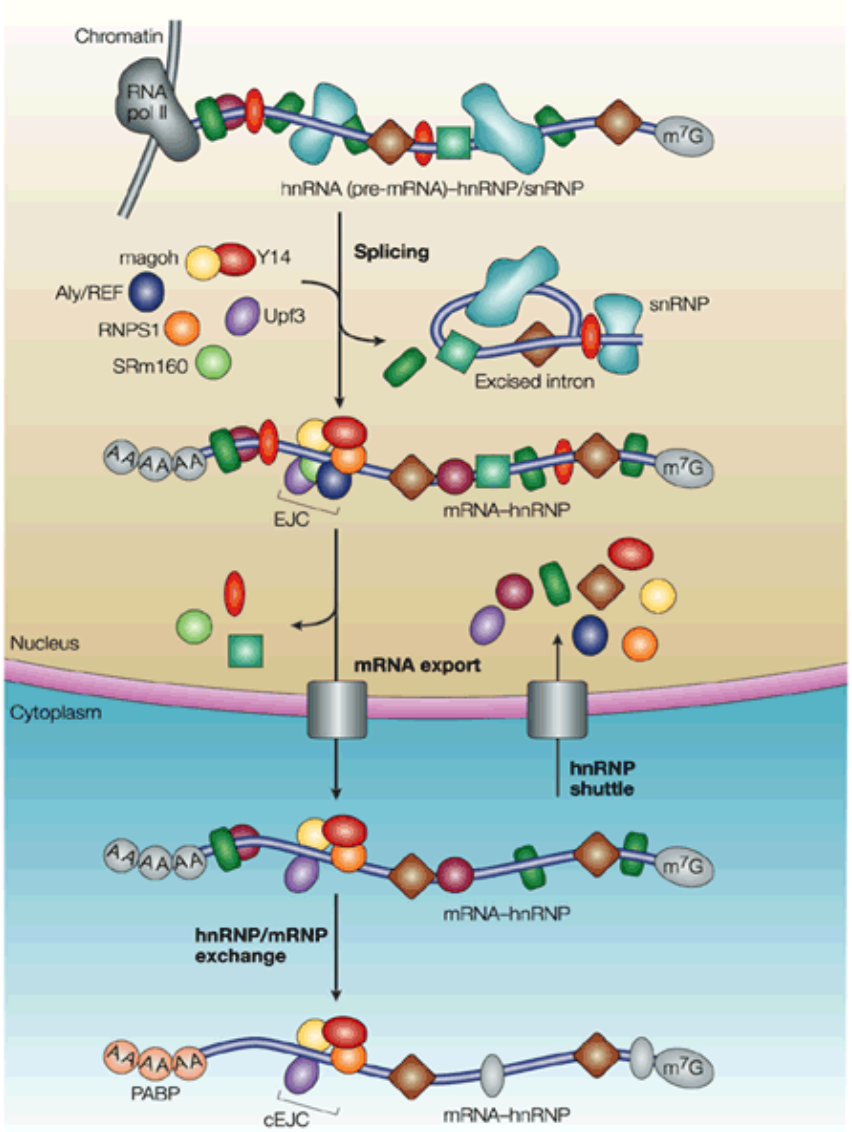
hnRNA modifications

- hnRNP-complexes forming
- adding cap to 5´- end
- polyadenylation of 3´- end
- splicing of hnRNA

hnRNP-complexes forming

- Proteins which specifically bind on hnRNA = **hnRNP-proteins**
- Proteins which specifically bind on small nuclear RNA (snRNA) = **snRNP-proteins**
- snRNP-proteins + snRNA = **snRNP-particles**
- hnRNA + hnRNP-proteins + snRNP-particles = **hnRNP-complex**
- snRNP-particles bind on introns and form **spliceosom**, which drive the hnRNA splicing
- hnRNP-proteins participate on transport of mRNA to cytoplasm

hnRNP-complexes forming

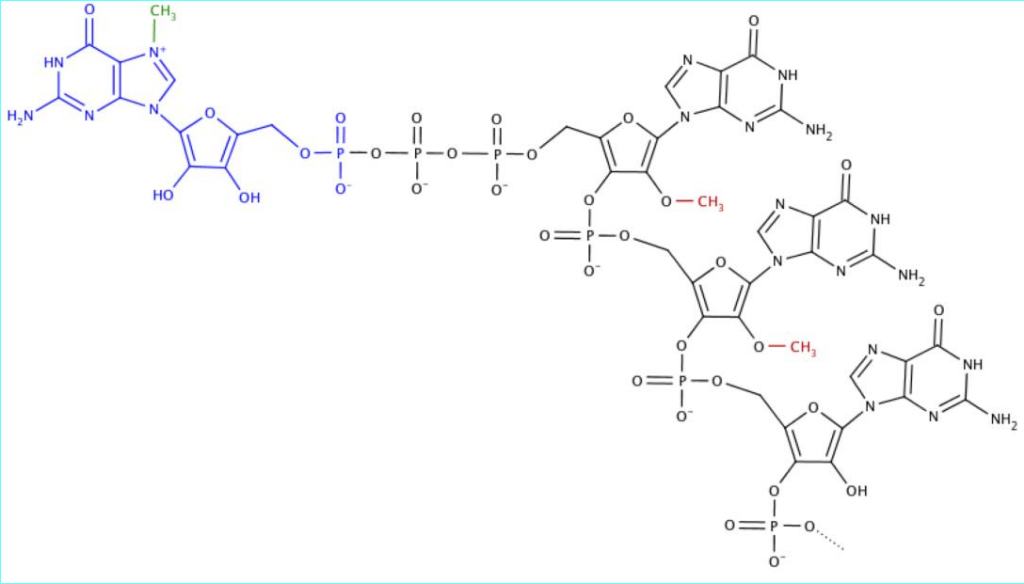


Nature Reviews Molecular Cell Biology **3**, 195-205 (March 2002)

Nature Reviews | Molecular Cell Biology

Adding cap to 5´- end

- Binding of 7-metylguanosine (**m^{7}G**) via three phosphate groups to 5´-end of hnRNA by 5´-5´ bound
- Last two 5´-end nucleotides could be aslo methylated
- m^{7}G plays important role during initiation of translation



pppNpRNA

Pi ↓ — RNA Triphosphatase (RTPase)

ppNpRNA

GTP, PPi ↓ — RNA Guanylyltransferase (GTase)

GpppNpRNA

SAM, SAH ↓ — RNA (guanine-N7)-Methyltransferase (N7MTase)

m^{7}GpppNpRNA (Cap-0 RNA)

SAM, SAH ↓ — RNA (nuceloside-2'-O-)-Methyltransferase (2'OMTase)

m^{7}GpppN$_m$pRNA (Cap-1 RNA)

DOI: 10.5772/56166

Polyadenylation of 3´- end

- Addition of of 50 – 250 adenosines to 3´-end of hnRNA = **poly(A) sequence**
- Catalyses by **poly(A)-polymerase**
- Poly(A)-polymerase is a subunit of complex, which binds on polyadenilation signal of hnRNA
- Poly(A) tail is important during transport of mRNA to cytoplasm and for its stabilisation

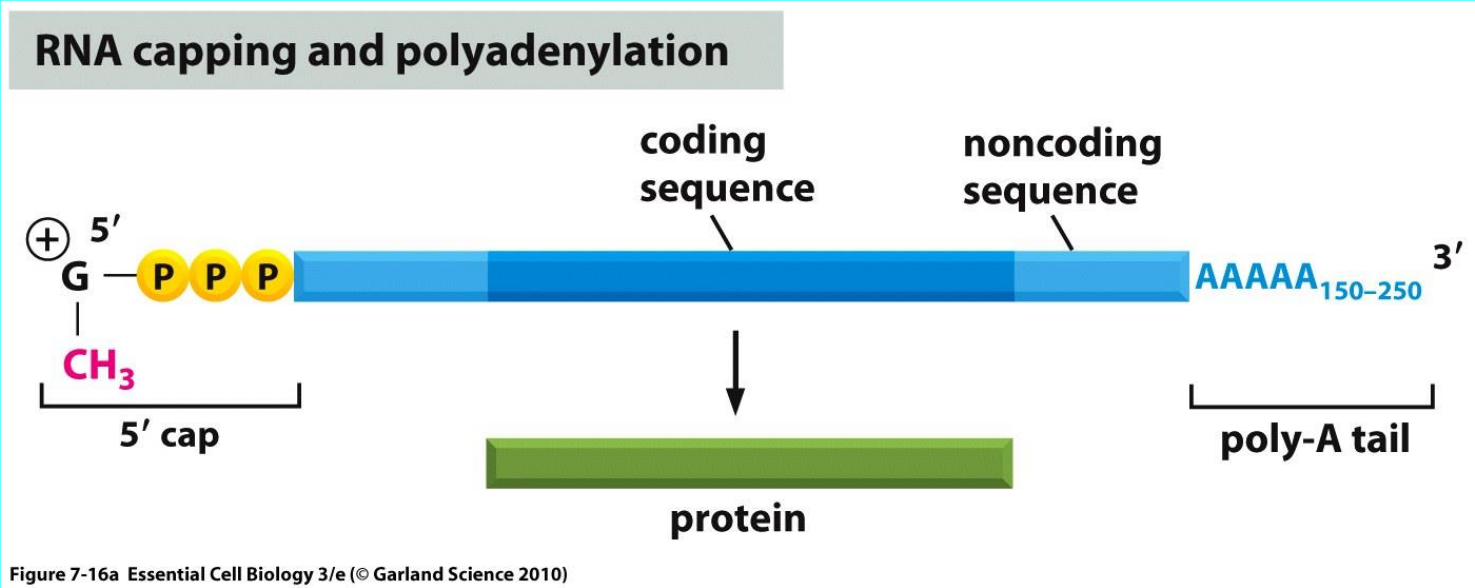
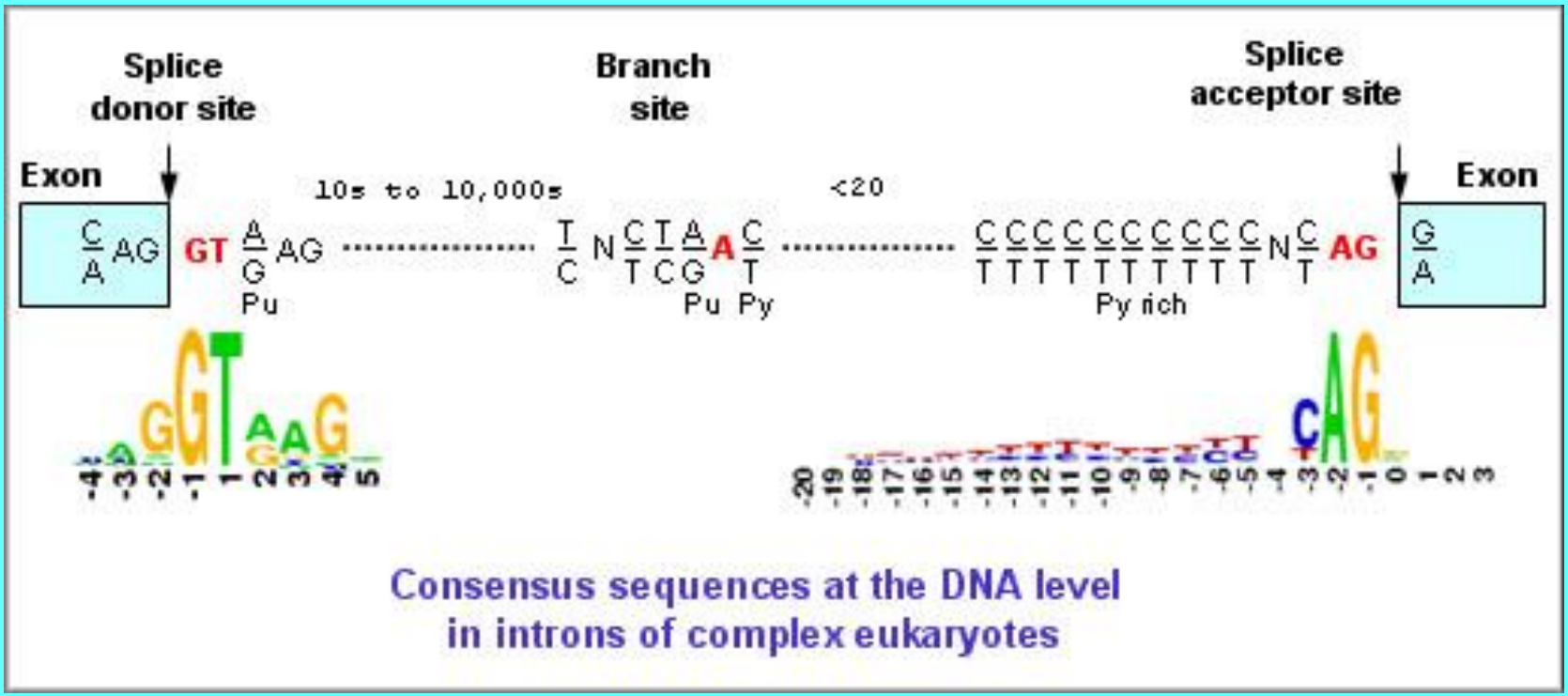


Figure 7-16a Essential Cell Biology 3/e (© Garland Science 2010)

Splicing of hnRNA

- **Introns** are cut out from hnRNA and mRNA is created
- Intron structure:
  - The rule of GU-AG
  - Branch site

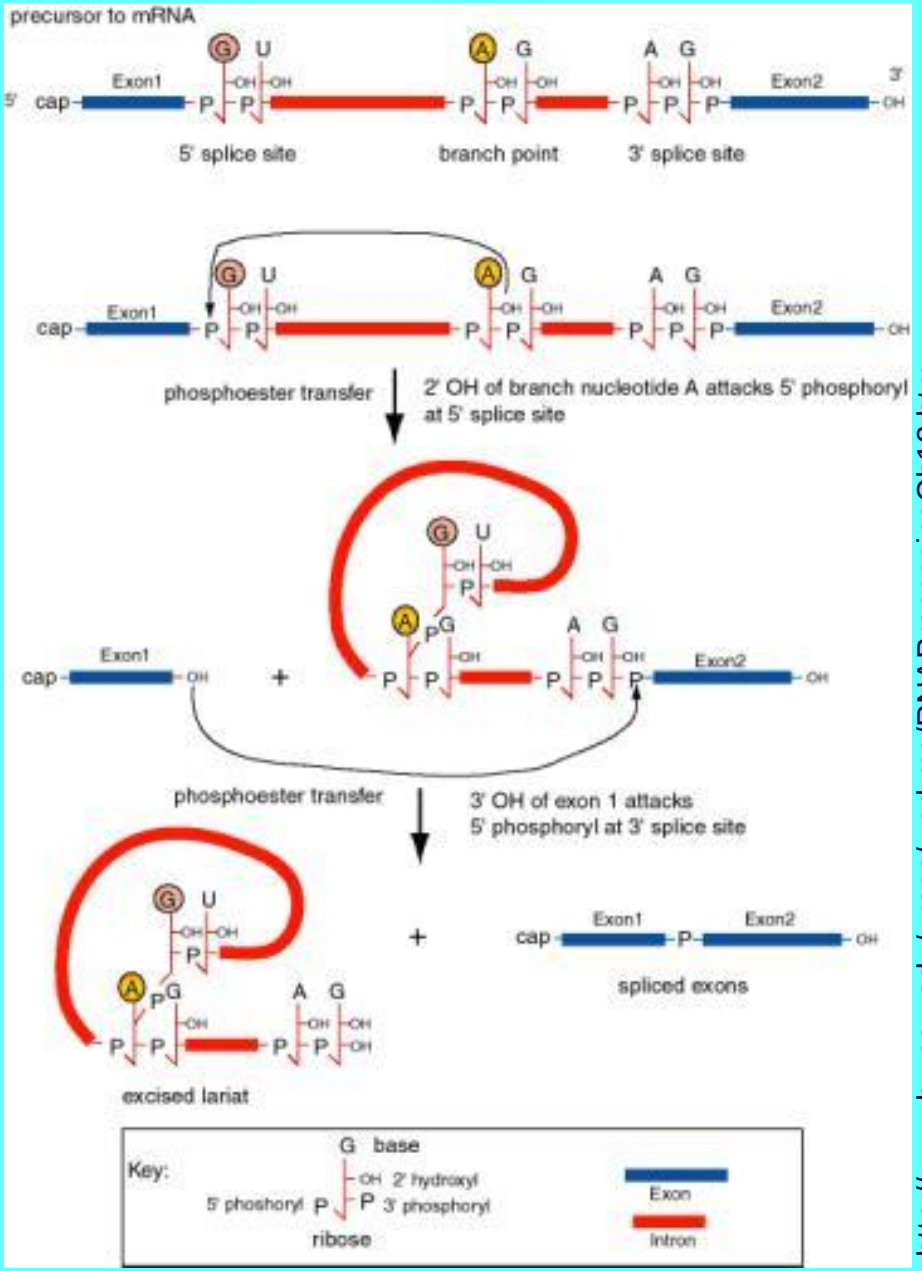
http://www.geneinfinity.org/sp/sp_coding.html



Splicing of hnRNA

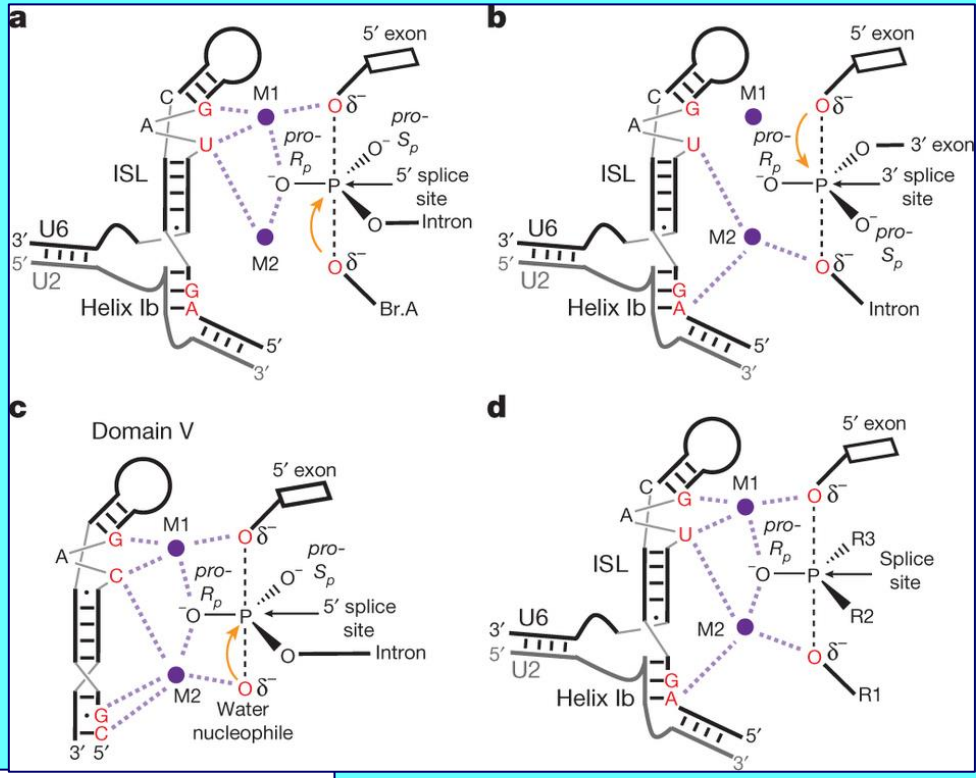
- Principle of splicing – **transesterification** – any energy from ATP or GTP is needed
- snRNA and snRNP-particles play the crucial role
- Reaction is catalysed probably by snRNA
- Intron is cut out in the form of **lariat intron RNA**

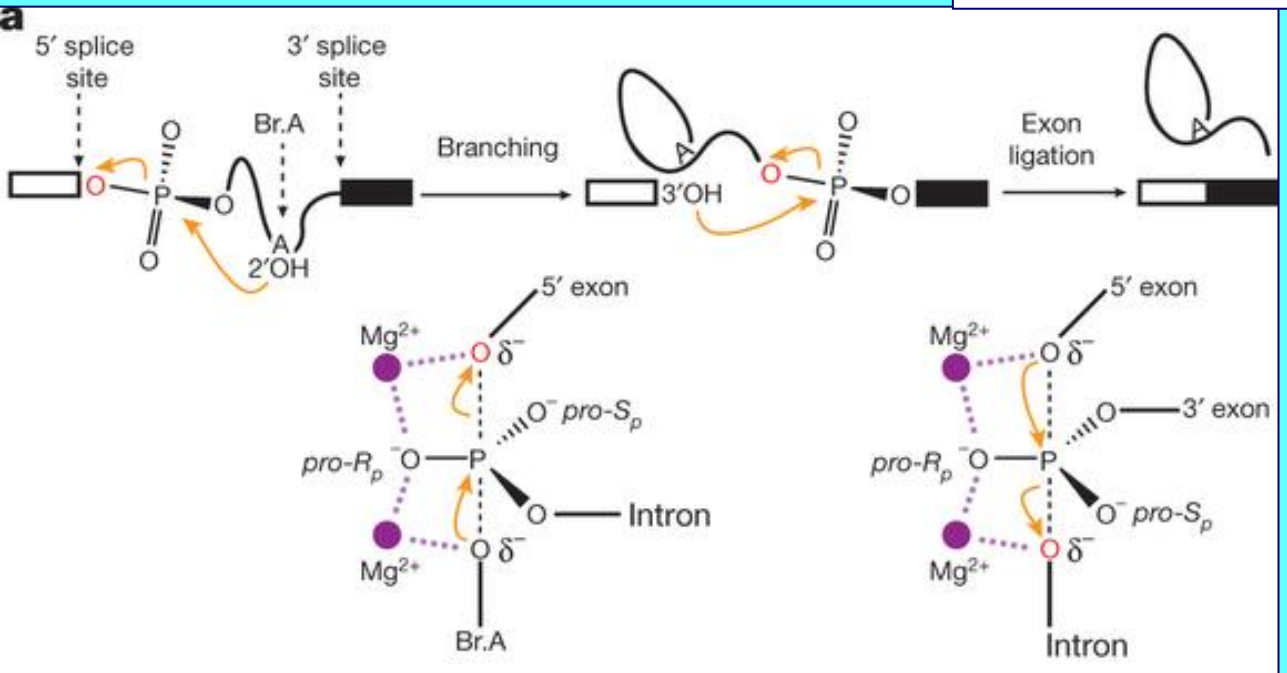
https://www.youtube.com/watch?v=YgmoHtLGb5c



http://www.bx.psu.edu/~ross/workmg/RNAProcessingCh12.htm

Role of M^{2+} ionts during splicing





Nature **503**, 229–234
(14 November 2013)
doi:10.1038/nature12734

Self-splicing

- **Autocatalytic process of introns and exons splicing.**
- **No proteins and enzymes are included in this process.**
- **Digestion and ligation of RNA substrate molecules during self-splicing is catalysed by ribosyme**

RNA editing

Posttranscription insertion or deletion of nucleotides in RNA strand or conversion of one base to another

Resulting in RNA transcript which sequence do not correspond to original sequence of DNA!!!

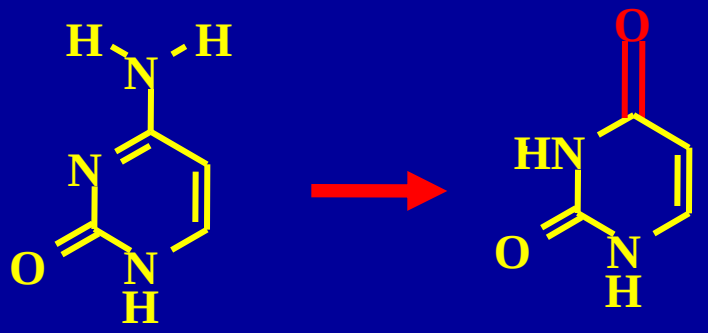
Structural genes undergoing of editation = kryptogenes

RNA editing was described in ´80 in *Trypanosoma*

The types of RNA editing

1) **Site specific deamination**
2) **gRNA-directed editing**

C → U deamination



cytosine **uracil**

Cytidin deaminase

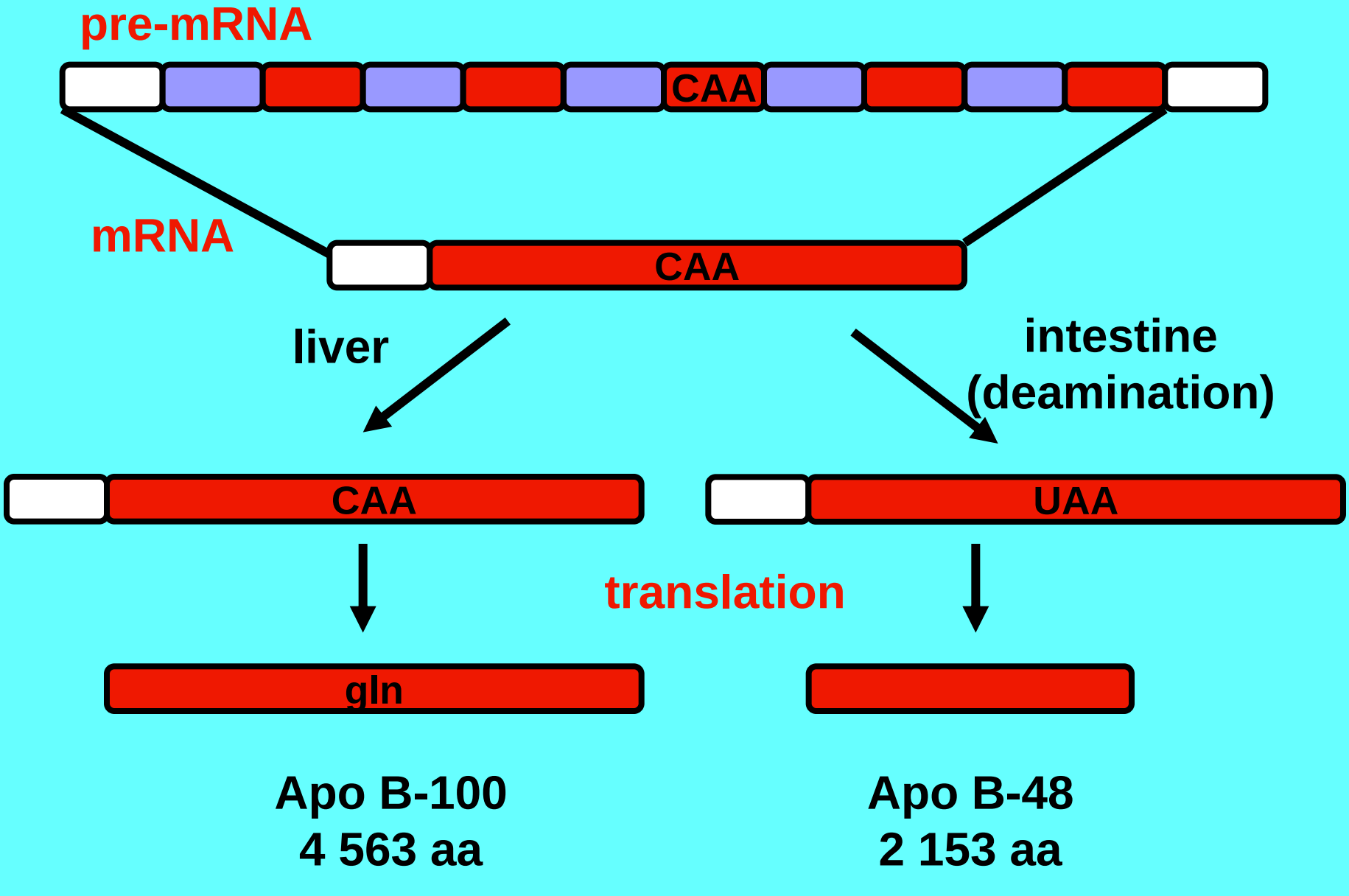
Effects of C → U deamination

- It proceed in specific mRNAs and only in certain tissues or cell types
- The process is regulated
- Two forms of apolipoprotein B arise

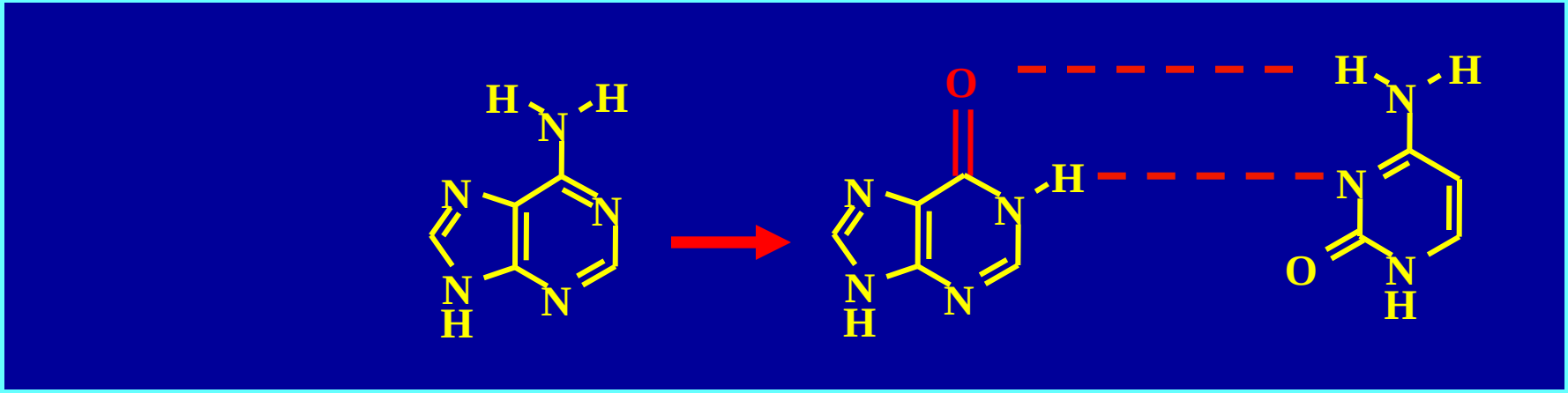
liver	intestine
Long	Short
Apo B-100	Apo B-48

Effect of the codon UAA formation

Editing of mRNA for apolipoprotein B



A → I deamination



adenine inosine cytosine

RNA specific adenosine deaminase (ADAR)

Effects of A → I deamination

- It proceed in ion channels of mammal brain
- Single nucleotide change proceeds to exchange of one amino acid
- This change permeability of ion channel to Ca^{2+} ions

If the process is inhibited serious damages of brain tissue development are found

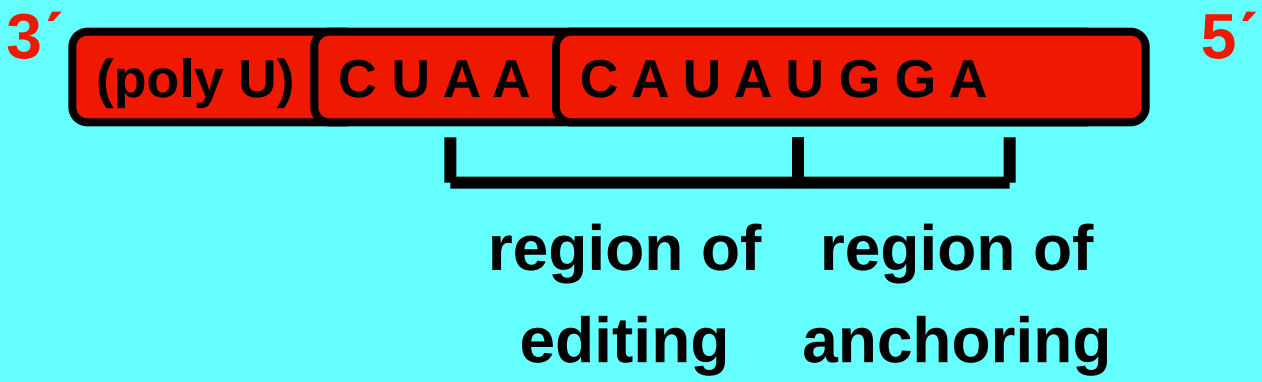
gRNA-directed editing

- **gRNA (guide RNA) are 40 - 80 nucleotides long**
- **Described in the *coxII* gene in *Trypanosome***
- **They enable adding of U in specific region of the transcript**
- **The resulting mRNA molecules contain additionally huge segments (inserts) which consist of U and opposite miss several U from original (maternal) DNA strand**
- **The inserts are such huge that finally up to 50% of edited mRNAs have post-transcriptionally added U**
- **The gRNAs joint to mRNA, enable their digestion, adding missing nucleotides and again ligation of digested segments**

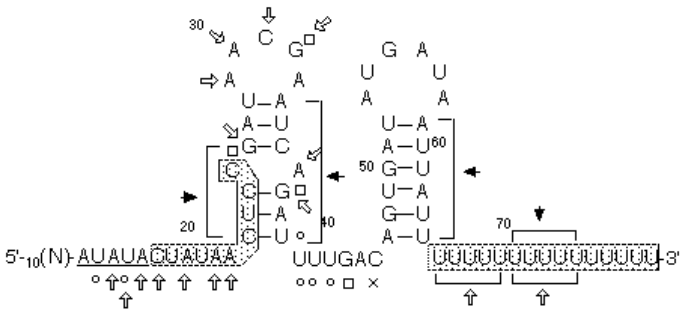
gRNA structure

Each gRNA has three regions

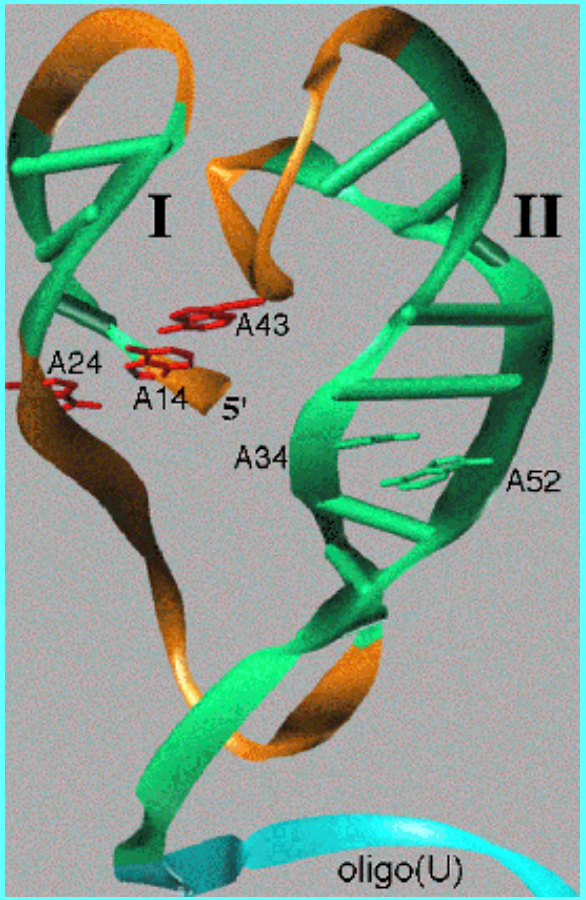
1) **The first, at the 5´- end (anchor) enables anchoring of gRNA to region of mRNA editing**
2) **The second direct which nucleotides will be added to edited sequence**
3) **The third, at the 3´- end is the polyU**



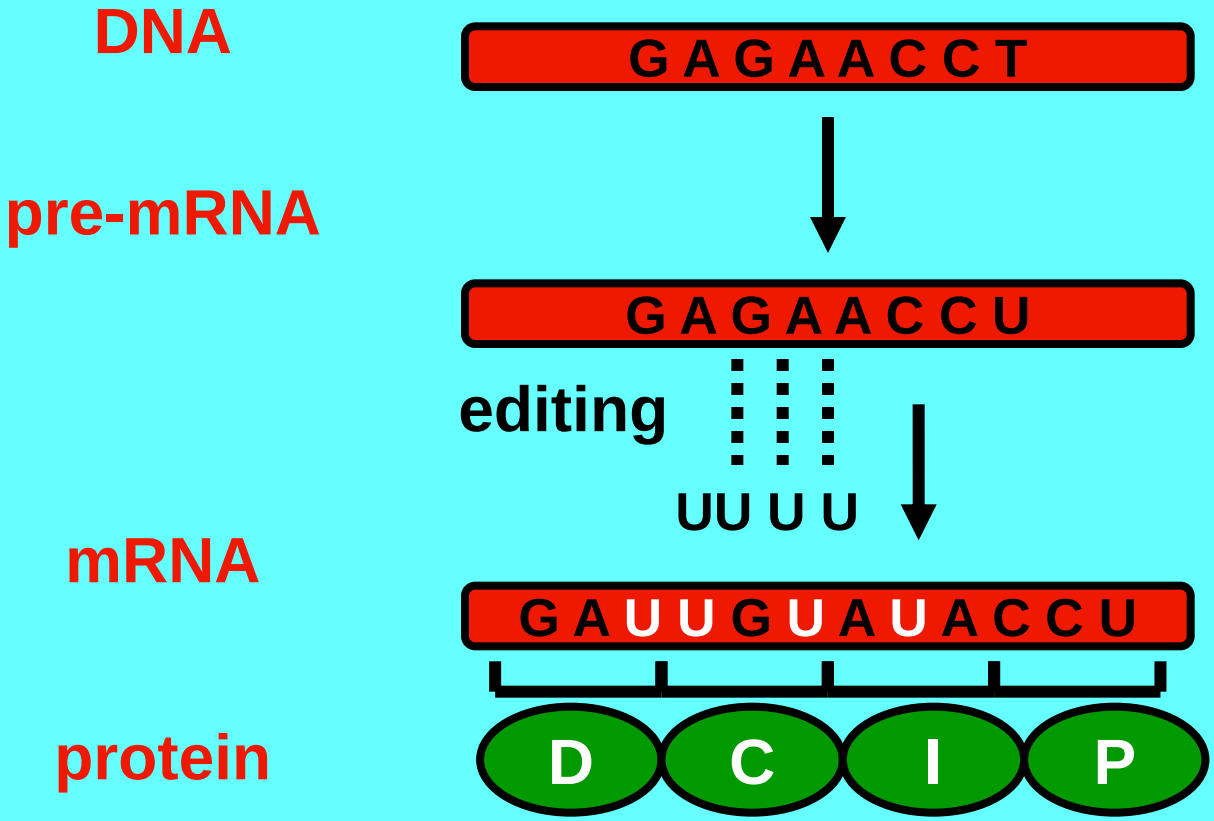
Example of gRNA-directed editing



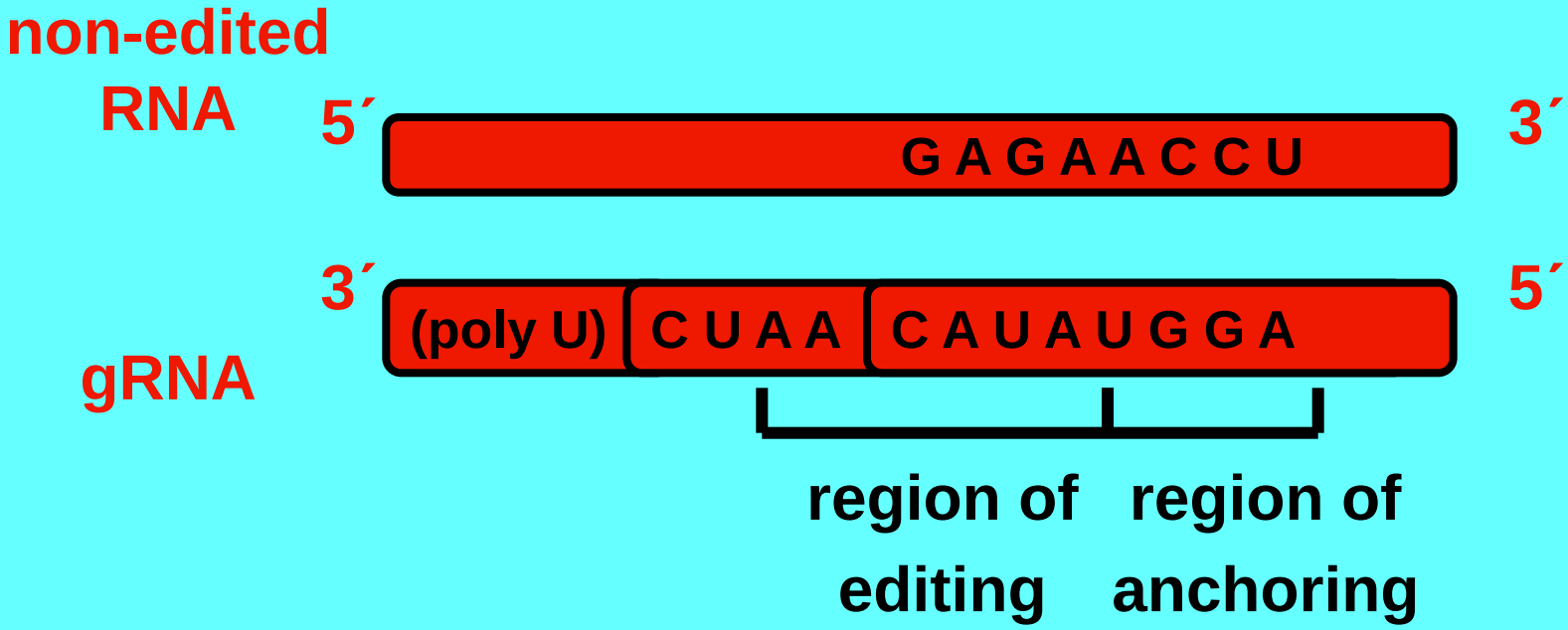
open square - kethoxal
filled circle - DMS
filled square - DEPC
open circle - CMCT
filled arrow + bracket - cobra venom nuclease
open arrow + bracket - T1, T2, S1 nucleases (i.e., probes for single-stranded regions)
x - frequent reverse transcriptase termination site in untreated control
boxed bases - anchor sequence and U-tail



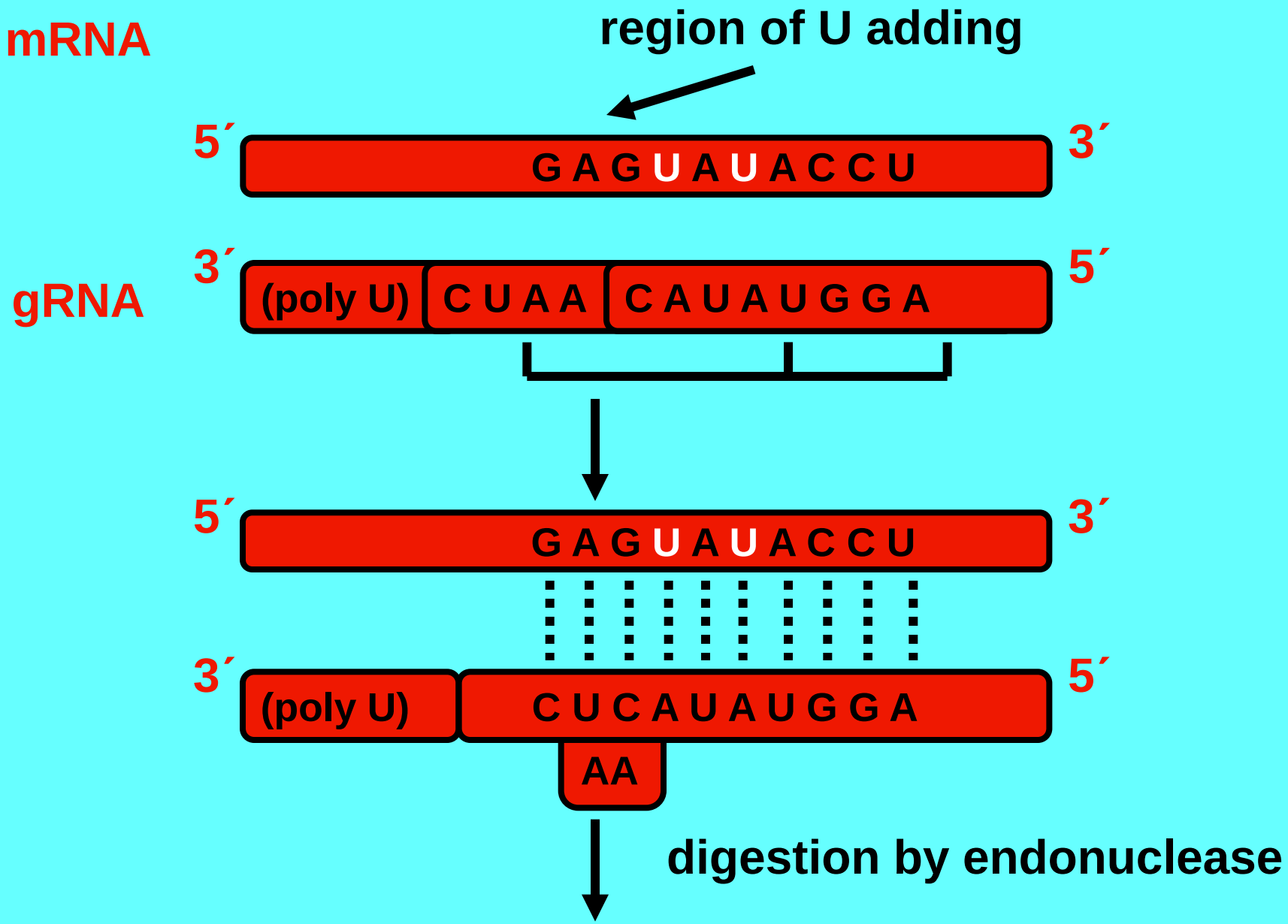
Position of four U nucleotides added to pre-mRNA of the coxII gene



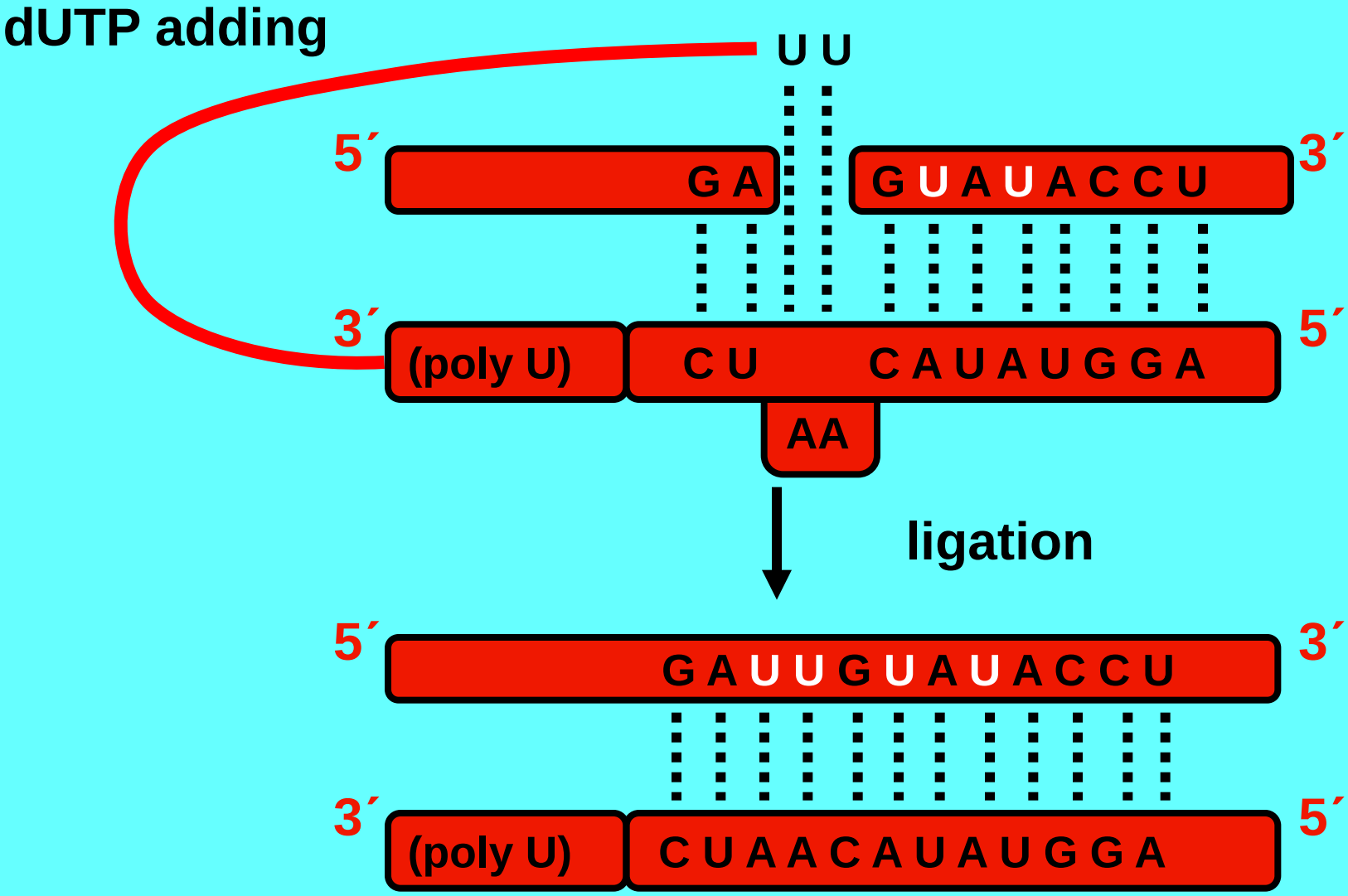
gRNA sequence and maternal non-edited mRNA



Process of editing - I



Process of editing - II



Eukaryotic translation

Differences to prokaryotic translation

- It proceeds in 2-3 compartements, cytoplasm,
- mitochondria, and chloroplasts
- The first AA is not fMet, but Met, which binds to a specific initiator $tRNA_i^{Met}$, which recognize the AUG codon
- The number of initiation factors which are necessary to beginning of translation is higher in eukaryotes
- The number of initiation factors which are necessary to beginning of translation is higher in eukaryotes

http://www.ncbi.nlm.nih.gov Taxonomy Genetic Codes

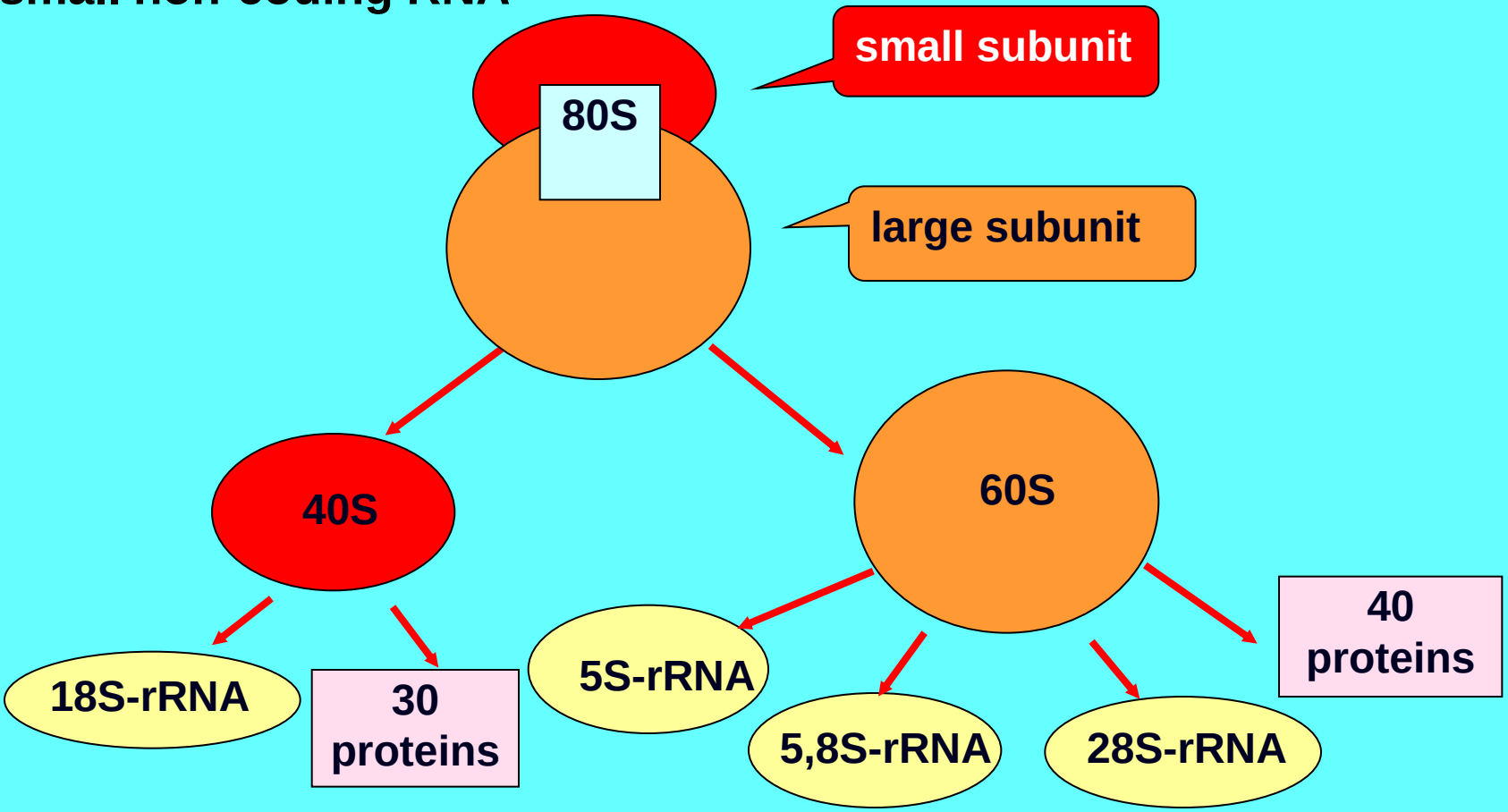
Course of translation

- similar as a translation in prokaryotes
- initiation, elongation, termination
- Particular complexes are more complicated
- More of translation factors
- Genetic code of mammalian mitochondria has different meaning of some codons, 22 tRNA
- Eukaryotic cell possesses 45 tRNA with different anticodons
- Speed of translation - 1-20 AA/s, depends on species and enviroment

The cytoplasmic ribosomes

Formation of ribosome structure involves also

- **150 non-ribosomal proteins**
- **100 small non-coding RNA**



Ferreira-Cerca, S. et al. (2007): Analysis of the In Vivo Assembly Pathway of Eukaryotic 40S Ribosomal Proteins, Molecular Cell 28, 446-457, November 2007

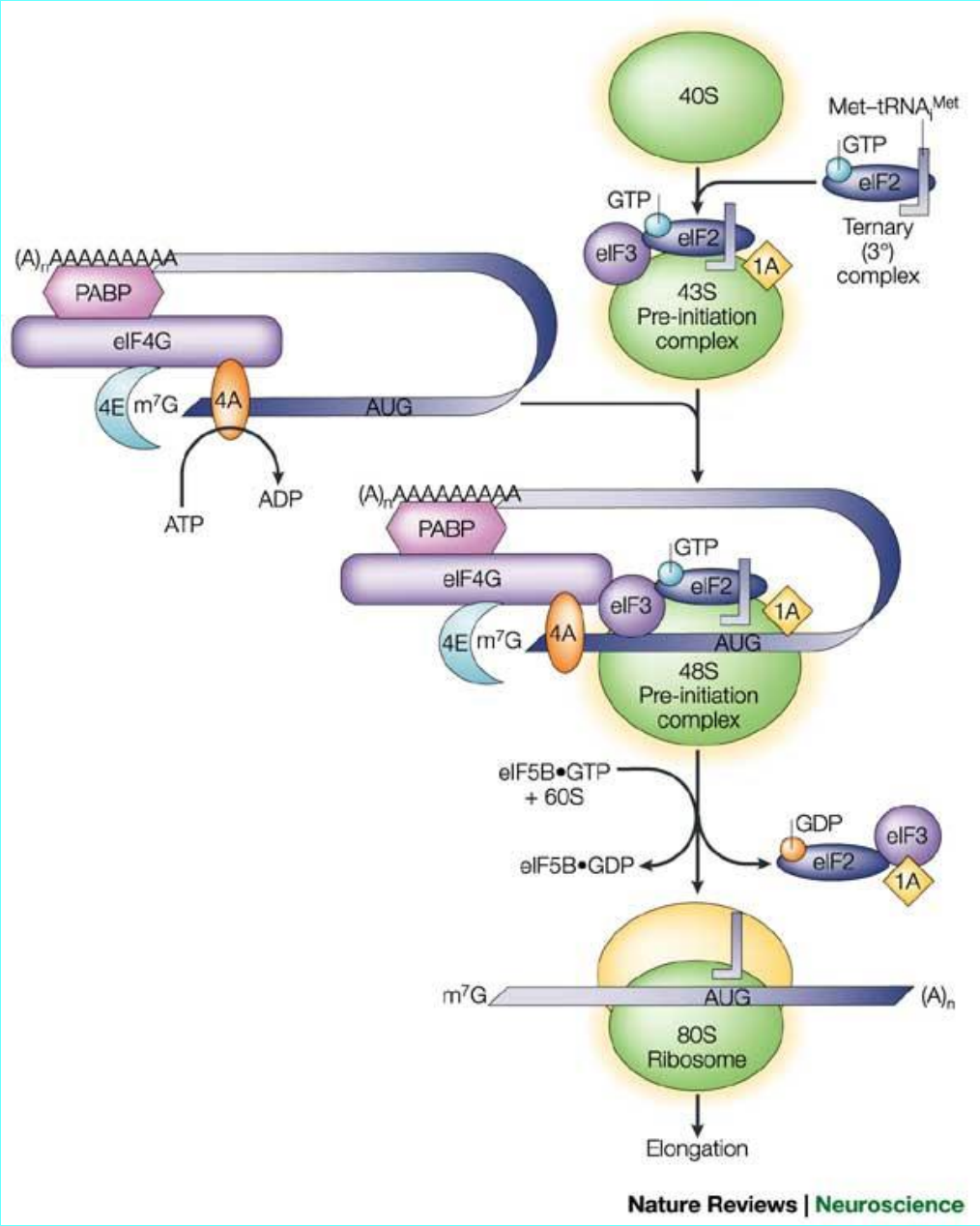
Free and bound ribosomes

- **Free ribosomes occur in cytoplasm**
  - **synthesis of intracellular proteins**

- **the rest is bounded to the endoplasmic reticulum**
  - **rough ER = covered by ribosomes**
  - **smooth ER = without ribosomes**
  - **synthesis of extracellular proteins**

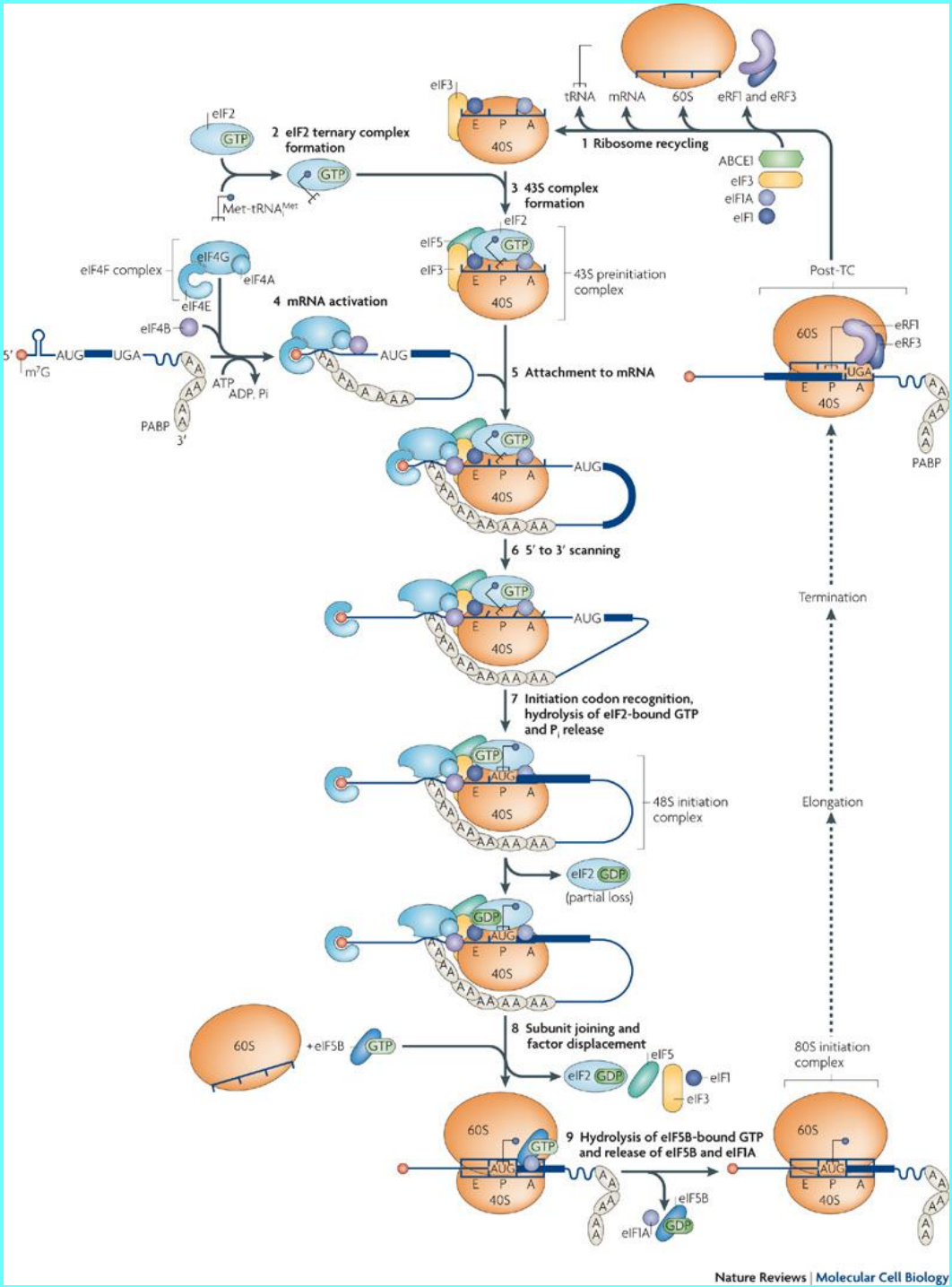
Initiation of translation

- 40S subunit with bound $tRNA_i^{Met}$ in P-site and initiation factors recognise m^7G cap of mRNA
- Subsequently, this complex moves to 3´-end until finds the initiation codon AUG
- Large 60S subunit binds to 40S subunit using the energy from hydrolysis of GTP



Nature Reviews Neuroscience **5**, 931-942 (December 2004)
doi:10.1038/nrn1557

Initiation of translation



Nature Reviews Molecular Cell Biology **11**, 113-127 (February 2010)
doi:10.1038/nrm2838

Initiation of translation – text to previous picture

The canonical pathway of eukaryotic translation initiation is divided into eight stages (**2**–**9**). These stages follow the recycling of post-termination complexes (post-TCs; **1**) to yield separated 40S and 60S ribosomal subunits, and result in the formation of an 80S ribosomal initiation complex, in which Met-tRNA$^{Met}_{i}$ is base paired with the initiation codon in the ribosomal P-site and which is competent to start the translation elongation stage. These stages are: eukaryotic initiation factor 2 (eIF2)–GTP–Met-tRNA$^{Met}_{i}$ ternary complex formation (**2**); formation of a 43S preinitiation complex comprising a 40S subunit, eIF1, eIF1A, eIF3, eIF2–GTP–Met-tRNA$^{Met}_{i}$ and probably eIF5 (**3**); mRNA activation, during which the mRNA cap-proximal region is unwound in an ATP-dependent manner by eIF4F with eIF4B (**4**); attachment of the 43S complex to this mRNA region (**5**); scanning of the 5′ UTR in a 5′ to 3′ direction by 43S complexes (**6**); recognition of the initiation codon and 48S initiation complex formation, which switches the scanning complex to a 'closed' conformation and leads to displacement of eIF1 to allow eIF5-mediated hydrolysis of eIF2-bound GTP and P_i release (**7**); joining of 60S subunits to 48S complexes and concomitant displacement of eIF2–GDP and other factors (eIF1, eIF3, eIF4B, eIF4F and eIF5) mediated by eIF5B (**8**); and GTP hydrolysis by eIF5B and release of eIF1A and GDP-bound eIF5B from assembled elongation-competent 80S ribosomes (**9**). Translation is a cyclical process, in which termination follows elongation and leads to recycling (**1**), which generates separated ribosomal subunits. The model omits potential 'closed loop' interactions involving poly(A)-binding protein (PABP), eukaryotic release factor 3 (eRF3) and eIF4F during recycling (see Supplementary information S5 (box)), and the recycling of eIF2–GDP by eIF2B. Whether eRF3 is still present on ribosomes at the recycling stage is unknown.

Termination of translation

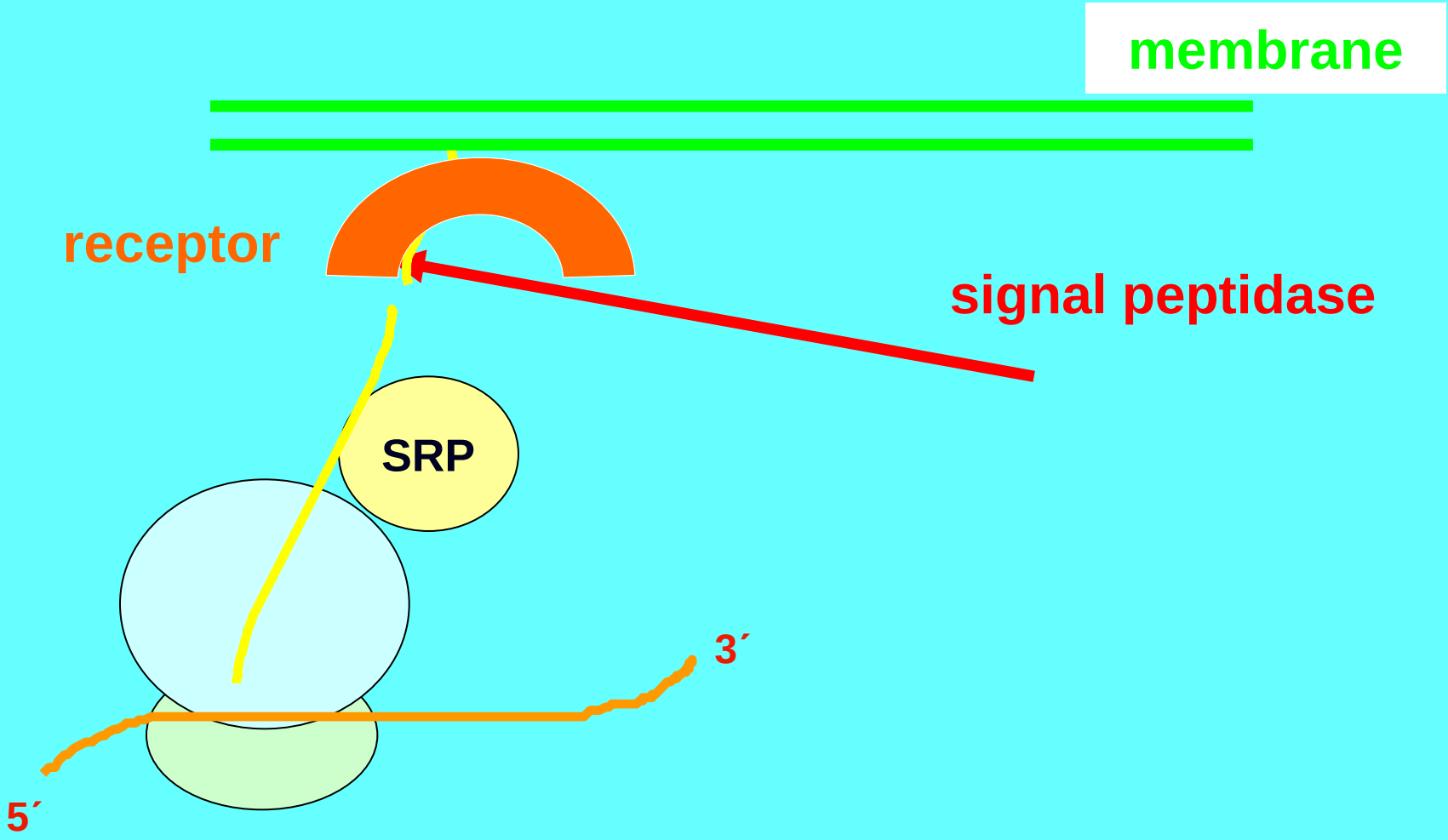
- **Only one termination factor = eRF**
- **Disociation of ribosome from mRNA needs the energy from GTP**

https://www.youtube.com/watch?v=qIwrhUrvX-k

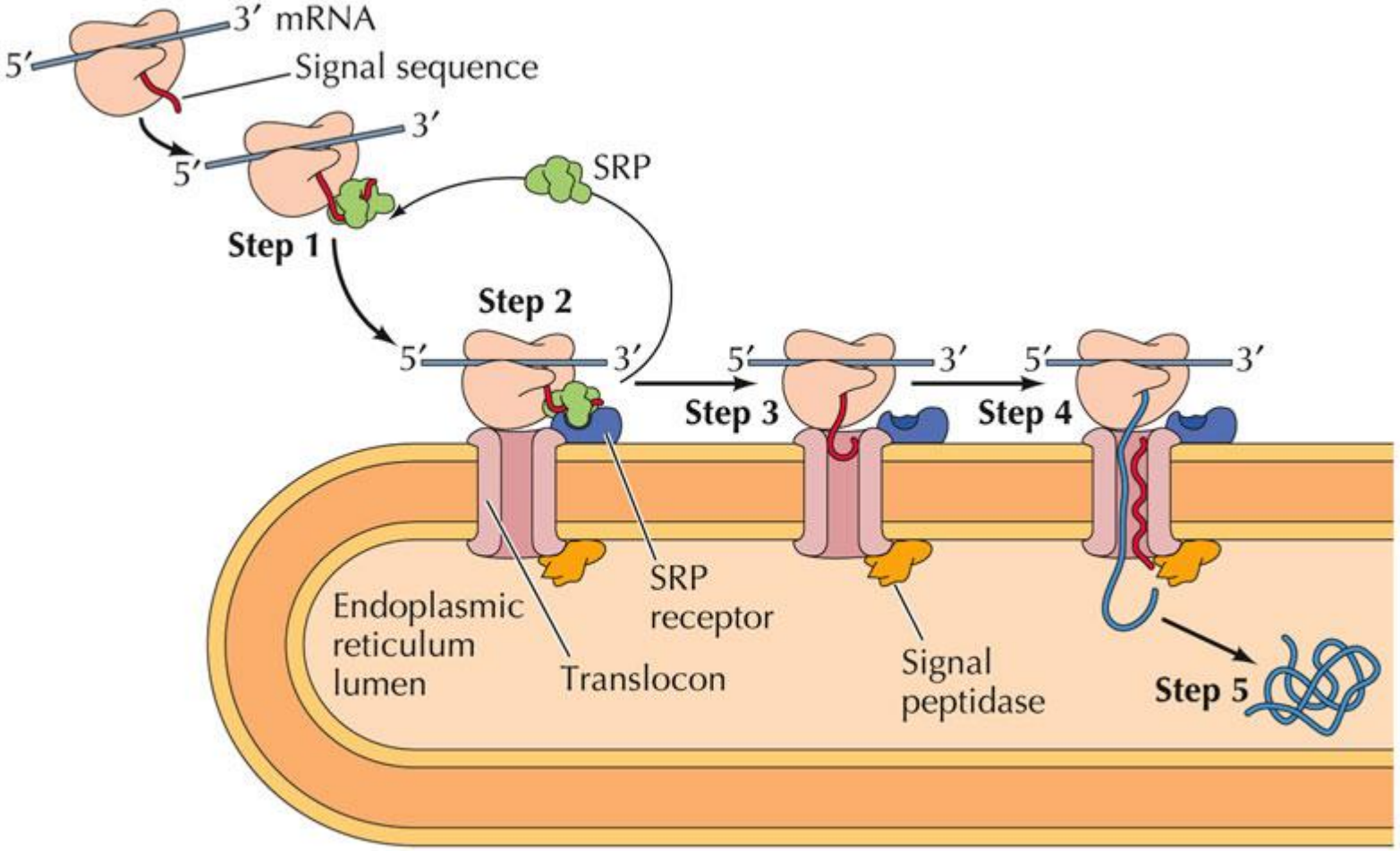
Extracellular end membrane proteins

- all extracellular and membrane proteins have on their N-end so named signal peptide 15-25 AA long
- signal peptide joints the proteins to signal recognition particle (SRP)
- SRP stops translation on ribosome
- binding of SRP to membrane receptor results in removing the signal peptide signal peptidase, and translation starts again

Extracellular end membrane proteins

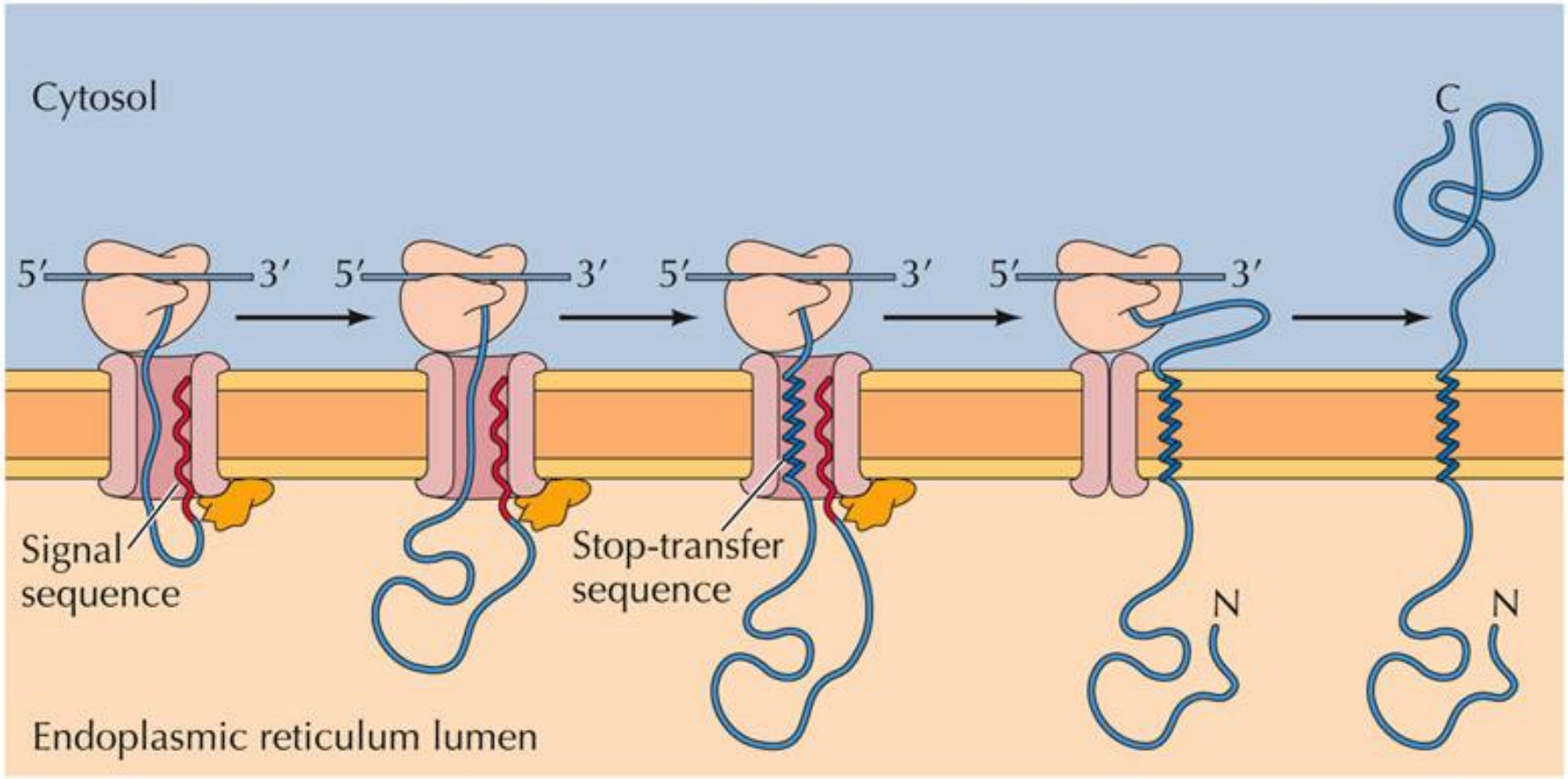


Translocation of extracellular proteins



THE CELL, Fourth Edition, Figure 10.8 © 2006 ASM Press and Sinauer Associates, Inc.

Formation of membrane bound proteins



THE CELL, Fourth Edition, Figure 10.12 © 2006 ASM Press and Sinauer Associates, Inc.

The structure of ribosome is still under intensive research

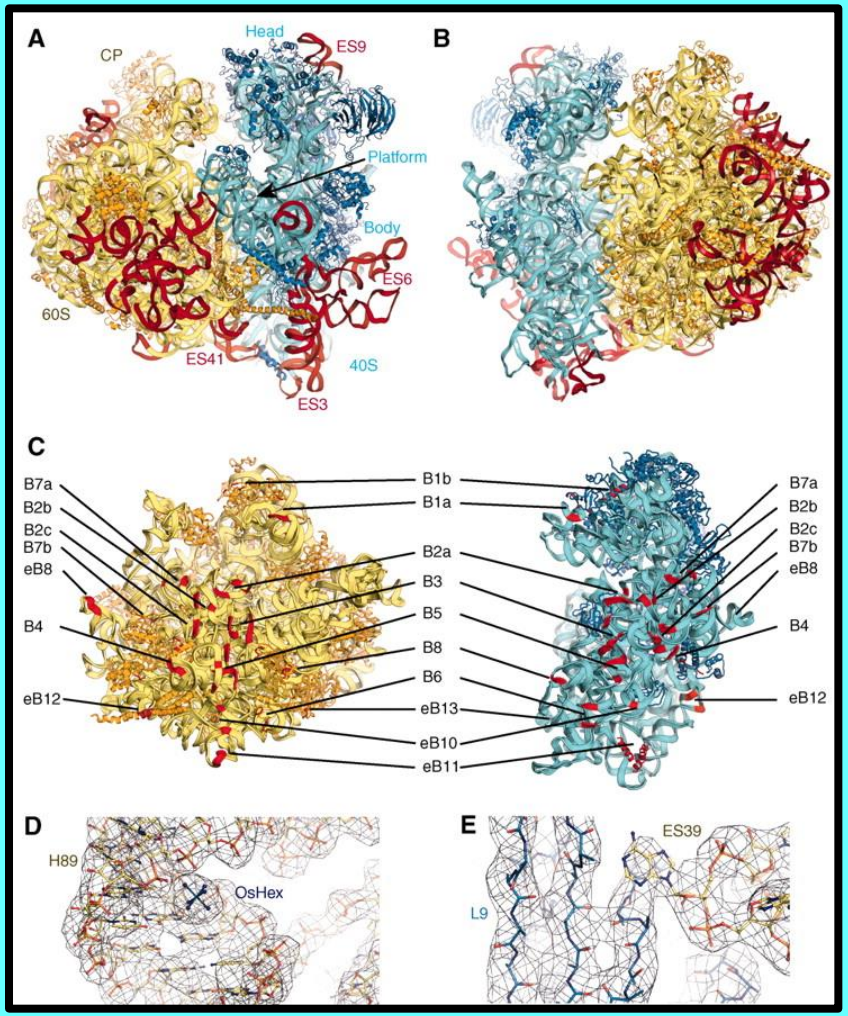
November 2010

Crystal structure of the eukaryotic ribosomes was described in resolution 4.15 Å

The most interesting thing was founding that both unit of ribosome fit as ratchets and during translation turn around themselves

Overall view of the x-ray structure

View from E-site



View from A-site

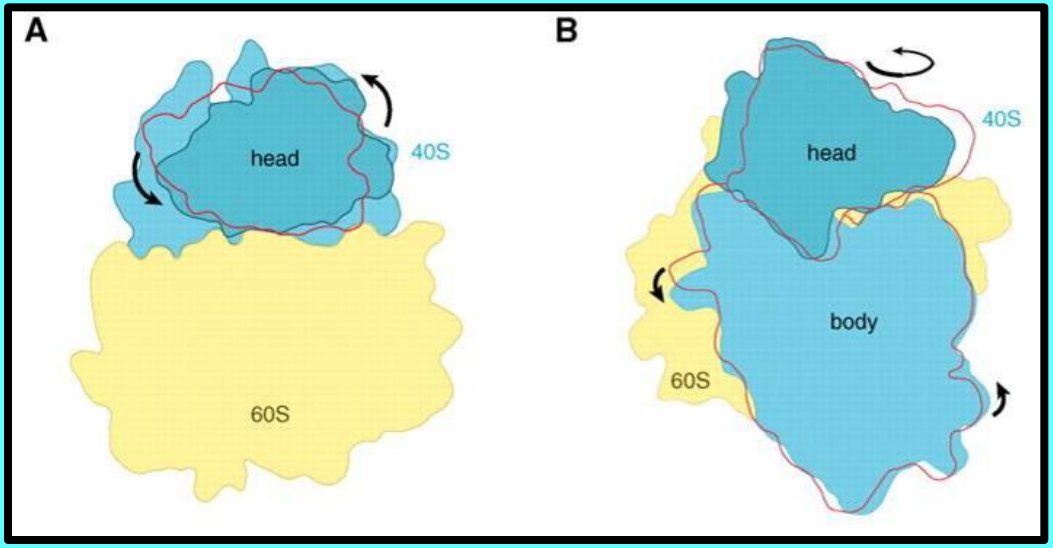
A Ben-Shem et al. Science 2010;330:1203-1209

Science AAAS

View to ratcheted state

top view

View from side of 40S

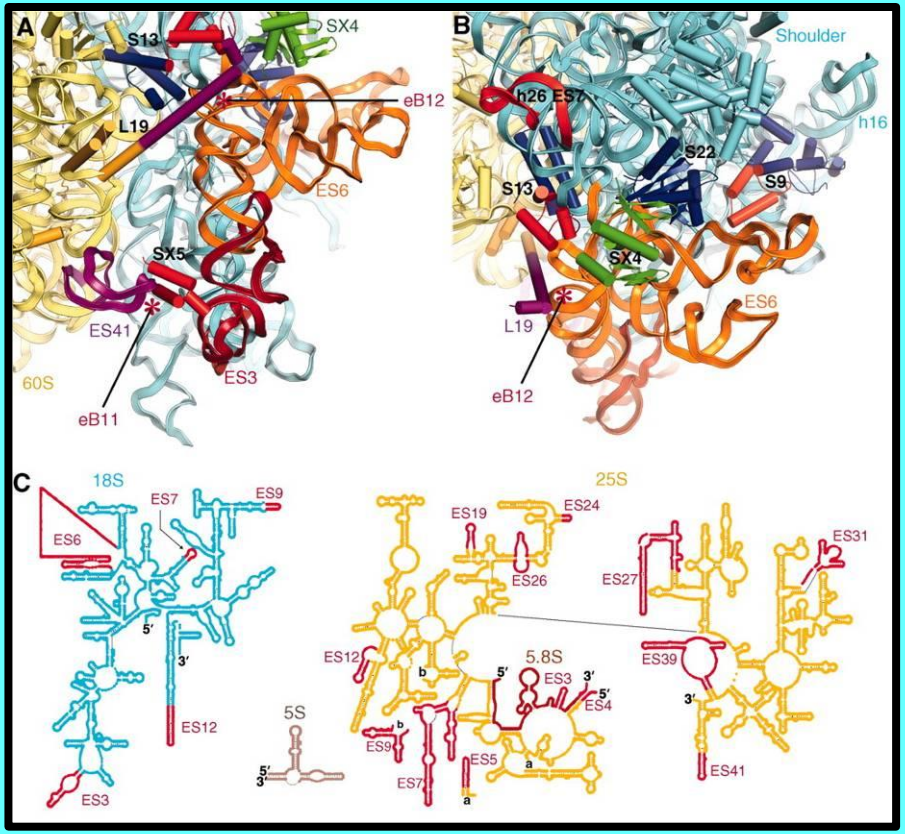


A Ben-Shem et al. Science 2010;330:1203-1209

Science AAAS

Network of interactions formed by eukaryote-specific elements

18S rRNA (blue)



5S rRNA (magenta)

25S rRNA (yellow)

5,8S (red)

Ben-Shem A. et al. Science 2010;330:1203-1209

Science AAAS

May 2011

The ribosome controls movement of tRNA and mRNA, structures described in resolution ~ 3.2 Å.

The structures help to explain how the ratchet-like motion of the two ribosomal subunits contributes to the mechanisms of translocation, termination, and ribosome recycling.

Dunkle et al. Science 2011;332:981-984

November 2011

Crystal structure of the large 60S eukaryotic ribosomes was described in resolution 3.5 Å

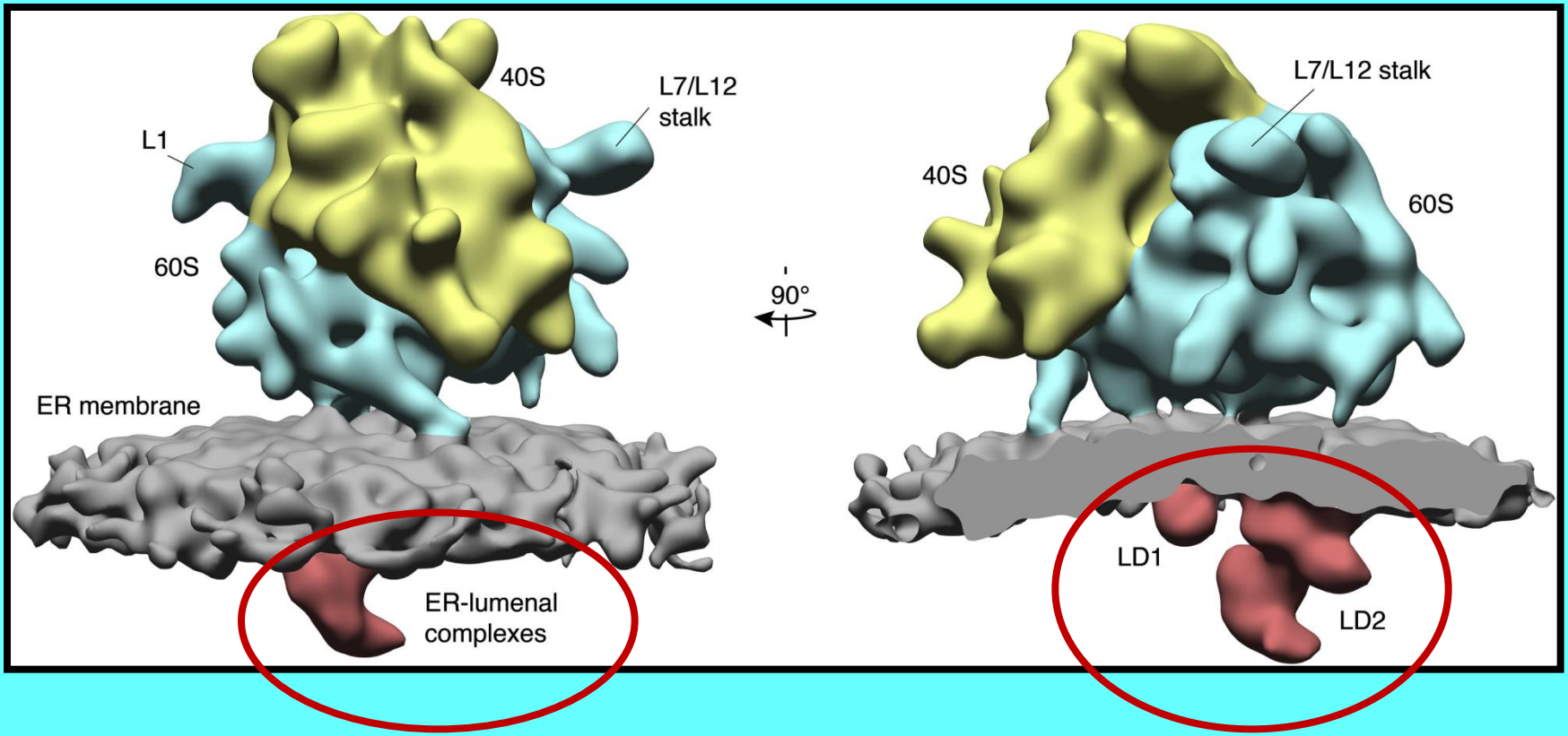
Klinge et al. Science 2011; 334 (6058): 941-948

September 2012

Crystal structure of the ribosomes bound endoplasmatic reticulum was described in resolution 31 Å

Pfeffer et al. (2012): Structure 20, 1508–1518

3D model of endoplasmatic reticulum bound ribosome



Pfeffer et al. (2012): Structure 20, 1508–1518